Identification of CLP-1 atypical calpain substrates in *Caenorhabditis elegans*

Laurence Newman

Department of Biochemistry, 2018

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of MSc by Research in the Faculty of Biomedical Sciences

Word Count: 26788

Authors Declaration

This project is my own work except where indicated. All text figures, tables, data, or results, which are not my own work are indicated and the sources acknowledged. In addition, I confirm that the hardcopy and the e-submission are identical.

SIGNED: …… DATE: ……..

Table of Contents

ACKNOWLEDGEMENTS

I would like to thank my father Richard, my older brother Fred, Marilyn, Phil and Rob for their support, they have all supported me both financially and emotionally, and I am eternally grateful to them all.

I wish to thank my supervisor Patty for all her research suggestions, guidance, support and time. I am very grateful to have had her as my supervisor.

I would also like to thank all the past and current members of Patty's lab that have helped me, especially Mengmeng. Finally, I'd like to thank Chris for his friendly support and letting me beat him at ping pong more than occasionally.

Abbreviations

^{137}Cs	Caesium-137
Ab	Antibody
API	Application programming interface
bp	Base pair
BSA	Bovine serum albumin
$CaCl_2$	Calcium Chloride
CAPN	Calpain (Mammalian)
cat. no:	Catalogue number
C. elegans	*Caenorhabditis elegans*
cDNA	Complementary deoxyribonucleic acid
co-IP	Co-immunoprecipitation
C-terminus	Carboxyl terminus
CysPC	Cysteine protease core domain
DSHB	Developmental studies hybridoma bank
dNTP	Deoxyribonucleotide triphosphates
DNA	Deoxyribonucleic acid
dTTP	Deoxythymidine triphosphate
DTT	Dithiothreitol
ECL	Enhanced Chemiluminescence
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol tetraacetic acid
EtOH	Ethanol
GTE	Glucose/ tris(hydroxymethyl)aminomethane)-base/ ethylenediaminetetraacetic acid
HCL	Hydrochloric acid
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HRP	Horseradish peroxidase
IgG	Immunoglobulin G
IFA	Intermediate filament A
IP	Immunoprecipitation
kb	Kilobase
kDa	kilo Dalton
KCl	Potassium Chloride
KH_2PO_4	Potassium Dihydrogen Phosphate
$K_2H\ PO_4$	Dipotassium hydrogen Phosphate
LB	Luria-Bertani medium
LC-MSMS	Liquid chromatography - tandem mass spectrometry
LGMD2a	Limb girdle muscular dystrophy type 2a
LP	Long-pass
MS	Mass spectrometry
$MgCl_2$	Magnesium Chloride
$MgSO_4$	Magnesium Sulphate
Min	Minute
MIT	Microtubule interacting and transport domain
mRFP	Monomeric red fluorescent protein
Na_2HPO_4	Disodium Hydrogen Phosphate
NaCl	Sodium Chloride

NaOH	Sodium Hydroxide
N/A	Not applicable
NGM	Nematode growth medium
nM	nanomolar
OD	Optical Density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate Buffered Saline
PBS-T	Phosphate Buffered Saline 0.1% (volume/volume) Tween-20
PCR	Polymerase chain reaction
PoI	Protein of interest
REST	Representational state transfer
Rpm	Revolutions per minute
Krpm	Thousand revolutions per minute
SDS	Sodium Dodecyl Sulphate
TEMED	Tetramethylethylenediamine
TML	Large trans membrane domain
TRIS	tris(hydroxymethyl)aminomethane)
TAE	tris(hydroxymethyl)aminomethane)-acetate ethylenediaminetetraacetic acid
TBE	tris(hydroxymethyl)aminomethane)-borate ethylenediaminetetraacetic acid
TE	tris(hydroxymethyl)aminomethane)-base/ ethylenediaminetetraacetic acid
TRIS-base	tris(hydroxymethyl)aminomethane)
v/v	Volume/volume
w/v	Weight/volume

Genetic abbreviations

rol-6	Roller 6
clp	Calpain (*C. elegans)*
egl	egg laying defective
asp	aspartyl protease
unc	uncoordinated
lon	long
myo	myosin heavy chain structural genes
pat	paralysed arrest at two-fold
dpy	dumpy
bli	blistered

Abstract

Calpains are regulatory calcium ion activated cysteine proteases present in nearly all eukaryotes (1). Dysregulation of Ca^{2+} homeostasis, which results in aberrant calpain activation is implicated in contributing to age-related pathologies, such as Alzheimer's disease and sarcopenia (2,3). However, it has been difficult to identify substrates targeted by calpain activation, because cleavage specificity is not determined by primary sequence alone (4).

Calpains are categorised into typical and atypical based on their domain architecture. Current understanding of human calpain activity is based primarily on studies of 'typical' calpains, which contain EF hand domains; less is understood about the six 'atypical' calpains that lack this domain. The genome of the model organism *Caenorhabditis elegans* encodes only atypical calpain proteins. Earlier studies have shown that these proteins are differentially expressed in a variety of tissues (5). Overexpression of CLP-1 in body wall muscle was shown to cause a low penetrance paralysis accompanied by muscle degeneration in adults (5). The severity of paralysis increased when CLP-1 was overexpressed in *egl-19(gf)* mutants that produce elevated levels of intracellular calcium. The substrates targeted by CLP-1 are unknown.

To identify potential CLP-1 substrates, a proteomic screen was performed. A CLP-1::mRFP fusion protein was expressed in muscle and protein complexes co-purifying with CLP-1 after immunoprecipitation were identified by mass spectrometry. A bioinformatics tool WormGRAB was created to collect information on proteins identified by mass spectrometry. A subset of proteins associated with muscle integrity and normal locomotion

were identified from the proteomic screen as well as other potential substrates that had not been previously associated with maintenance of muscle.

Further analysis of the intracellular localisation of CLP-1 in body wall muscle was also pursued to understand how it might be targeted to substructures of the sarcomere. The involvement of calpains in heat stress was also investigated.

List of Figures

List of Tables

1 Introduction

1.1 Overview of Calpains

Calpains are a family of calcium activated intracellular non-lysosomal cysteine proteases. Calpains were first purified from rat tissue by Guroff in 1964, and the name calpain was proposed by Murachi et al. in 1981 in recognition of their similarity to the calcium binding proteins calmodulin and the cysteine protease papain from papaya (6,7). Whole genome sequencing projects have shown that genes encoding predicted calpain products are present in a wide and extensive range of eukaryotes and eubacteria (8). They are commonly classified into typical and atypical groups, depending on the presence or absence respectively; of calcium binding EF-hand domains (Figure 1) (1). Calpains are now defined by presence of a Cysteine Protease Core domain CysPC (also referred to as domain II), which contains the catalytic triad (Figure 1). These multi-domain proteins are further classified based on their domain composition and expression pattern.

Unlike strictly degradative proteases, such as lysosomal proteases, which digest substrates into non-functional polypeptides and amino acids, calpains cleave substrates in a limited manner to produce a functional product with modulated activity. For this reason, calpains are often referred to as modulatory or regulatory proteases (9). Calpains are implicated as participating in numerous cellular processes; however their precise role is unclear in these processes due to a lack of identified physiological substrates (10).

Calpains are potential therapeutic targets as they are associated with a number of degenerative diseases including Alzheimer's disease (3). Limb girdle muscular dystrophy type 2a (LGMD2a), which is also referred to as calpainopathy, is caused by a mutation in the human muscle specific calpain gene CAPN3 (11). There are 15 calpain genes present in humans: 9 typical and 6 atypical (9).

1.2 Calpain structure and classification

Calpains are commonly classified by their domain organisation into ‘typical’ and ‘atypical‘ calpains. Typical calpains, all of which have a cysteine protease core domain, are further defined by their structural similarity to vertebrate CAPN1 and CAPN2, each of which have a penta EF-hand (PEF) domain and C2-Like domain (Figure 2) (10). Atypical calpains lack the PEF domain and might also lack C2-L domains. Recent phylogenetic analyses of calpains have revealed that the PEF hand domain of typical calpains was an evolutionarily late addition (8). This suggests that the PEF hand domain might act in conjunction with the vertebrate calpain small subunit (CAPNS) to enable binding of the endogenous inhibitor calpastatin as an additional regulatory mechanism (8,12). Despite the greater abundance of genes encoding atypical calpains across eukaryotes, the majority of research has focused on typical calpain function and their regulation (10).

1.2.1 Phylogeny

Calpains are present in eukaryotes and bacteria, but are absent from archaea (8). Eukaryotic calpains are likely to have originated by horizontal gene transfer from bacteria (13). Whereas bacteria have been found to express at most one calpain gene, eukaryotes may express several calpain genes with different domain architectures. Calpains are present in the vast majority of eukaryotes leading to earlier hypothesis that calpains are essential genes in eukaryotes, but they have been found to be absent from several unicellular eukaryotes (8).

As shown in Figure 1, phylogenetic analysis of eukaryotes has revealed the existence of four ancestral domain architectures (CysPC, CysPc-C2L, TML-CysPc-C2l, MIT-CysPc-C2L) of eukaryote calpains (8). Typical calpains that have both C2-L and EF hand

domains have evolved relatively recently in eukaryotes, and are absent in bacteria, represent a minority subset of the calpain family (13). Typical calpains have evolved several regulatory mechanisms; EF-hand domains, small subunit interaction, calpastatin interaction (1). Although activity of atypical calpains are likely to be primarily regulated by calcium ion levels, other regulatory mechanisms may exist amongst the variety of domains present in atypical calpains. The different atypical calpain subfamilies are described in Section 1.2.

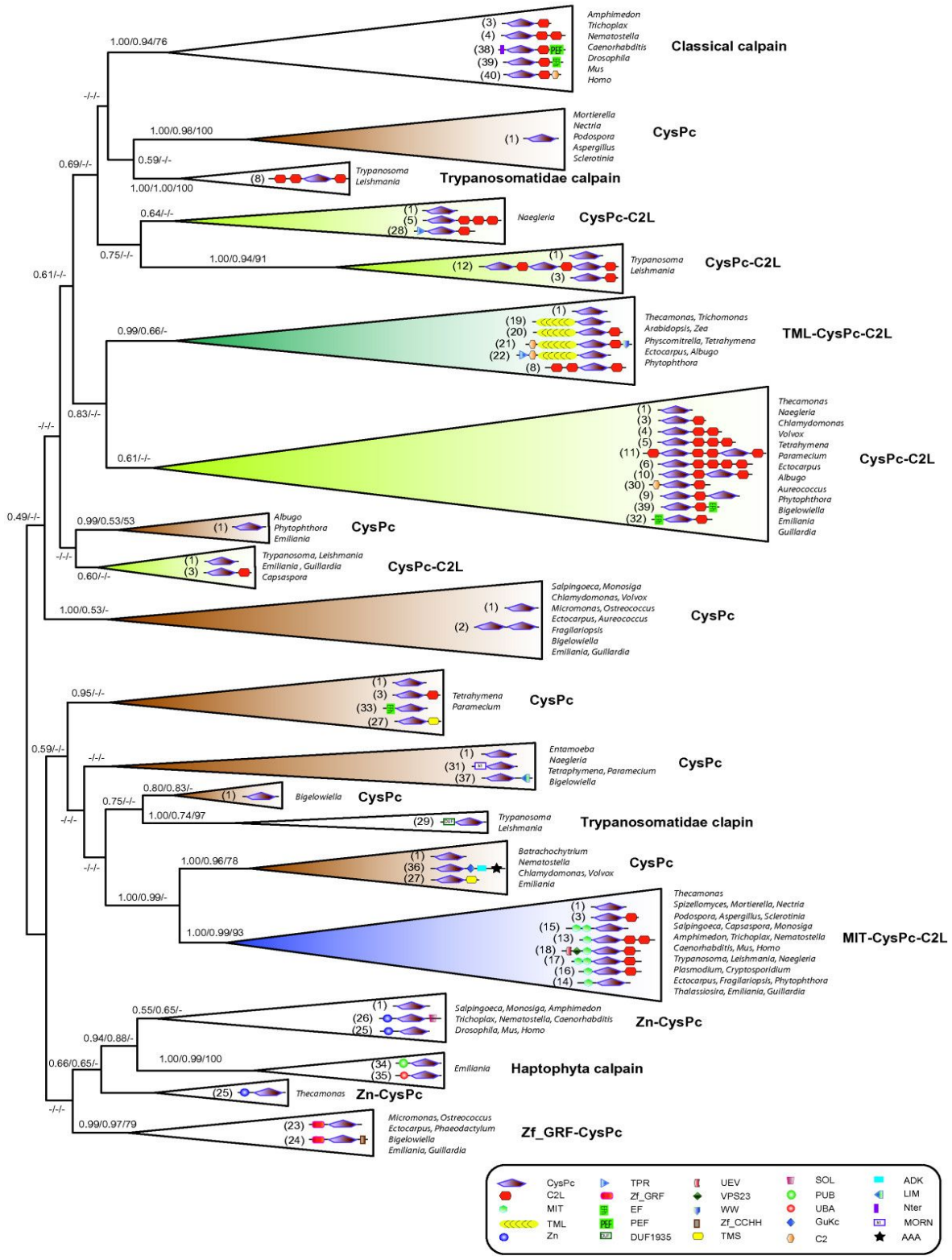


Figure 1. Eukaryotic calpain phylogeny.

The clades representing four proposed ancestral domain architectures are color-coded (i.e. CysPc in dark brown, CysPc-C2L in blue, MIT-CysPc-C2L in light green and TML-CysPc-C2L in dark green).
Classical (Typical) calpains have evolved relatively recently. Adapted from Zhao et al. 2012 (8).

1.2.2 Structure and characterisation

The characterisation of the calpain family has primarily focused on the typical calpains μ-calpain and m-calpain (known today as CAPN1 and CAPN2). These proteins are named according to the micro and milli molar concentrations of Ca^{2+}, respectively, at which they become proteolytically activated in vitro (14). CAPN1 and CAPN2 are expressed ubiquitously in mammals (15). Despite the difference in their Ca^{2+} requirements, CAPN1 and CAPN2 (Figure 2) share over 60% amino acid identity and are both inhibited by the only identified endogenous calpain inhibitor calpastatin (1,16). Although CAPN1 and CAPN2 were originally suspected to have similar substrate specificity, more recent studies have revealed differences in substrate preferences for each protease (17).

The typical calpains CAPN1 and CAPN2 each consist of a large catalytic subunit that associates with a common small regulatory subunit CAPNS. CAPNS is encoded by a separate gene. As shown in Figure 2, the large subunit ~80 kDa is organised into four domains (I-IV) and the small subunit ~28 kDa two domains (V-VI) (10). Domain boundaries were refined when the crystal structure of rat and human m-calpain was solved (18,19). Domain I consists of an α-helix that acts as an anchor between the protease core domain and the regulatory subunit domain VI. Domain I undergoes autolysis, which in turn lowers the concentration of Ca^{2+} required for proteolytic activity (Figure 3) (20). Autolysis is one of several ways in which the activity of the conventional calpains is regulated (18,21).

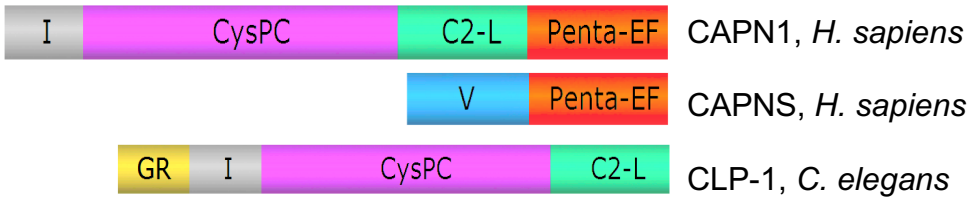


Figure 2. Human typical calpain CAPN1 and *C. elegans* atypical calpain CLP-1 proteins.

Domain architecture of the human typical CAPN1 and CAPNS (μ-calpain), and the *C. elegans* non-classical calpain CLP-1. CysPC domain is the proteolytic domain. C2-L refers to a calmodulin like domain. PEF domain contains 5 EF-hand motifs. GR refers to a glycine rich region in CLP-1.

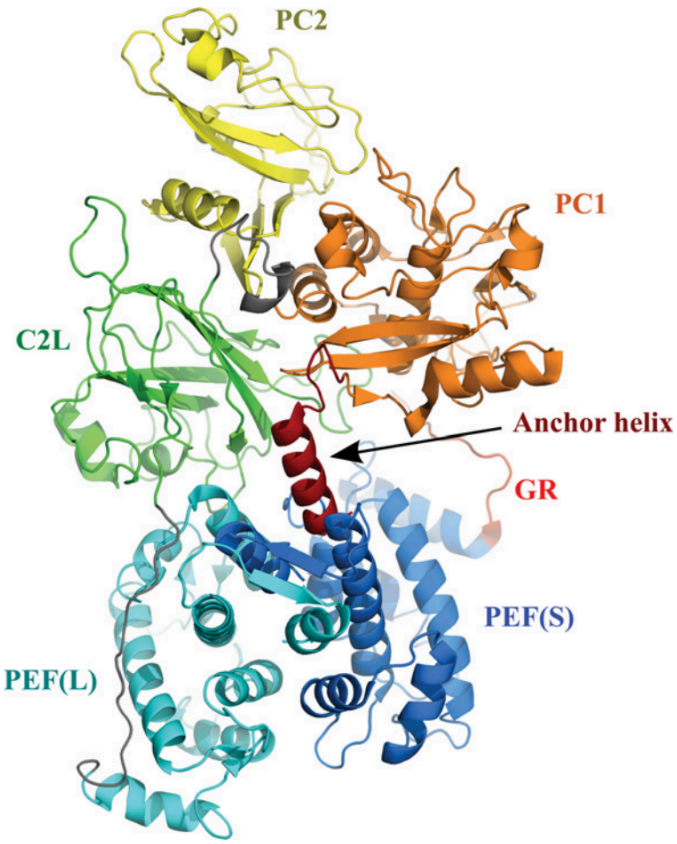


Figure 3. Anchor helix domain 1 undergoes autolysis allowing calpain activation.

Autolysis of domain 1 is hypothesised to promote PC1 and PC2 to undergo a conformational change activating the catalytic triad (20). Figure adapted from Campbell et al. 2012 (20).

Calpains are defined based on their amino acid sequence similarity to the protease core domain CysPC (domain II). CysPC contains an active site that houses the catalytic triad of cysteine, histidine and asparagine. The catalytic triad facilitates substrate proteolysis by covalent catalysis via nucleophilic substitution (22). The cysteine is activated by a nearby histidine, which is polarised after alignment with the asparagine. The crystal structure of Ca^{2+} free CAPN2 revealed that domain II consists of two subdomains PC1 and PC2 (Figure 3). PC1 contains the cysteine residue, which is 10.5 angstroms away from the histidine and asparagine of the catalytic triad. Binding of Ca^{2+} to the N-terminal anchor helix and CysPC causes a conformational change in PC1 and PC2, which was observed in the crystal structure of Ca^{2+} bound CAPN1 (see Figure 4). This conformational change brings the amino acids of the catalytic triad in close proximity to one another, causing the calpain to become proteolytically active (1,23,24).

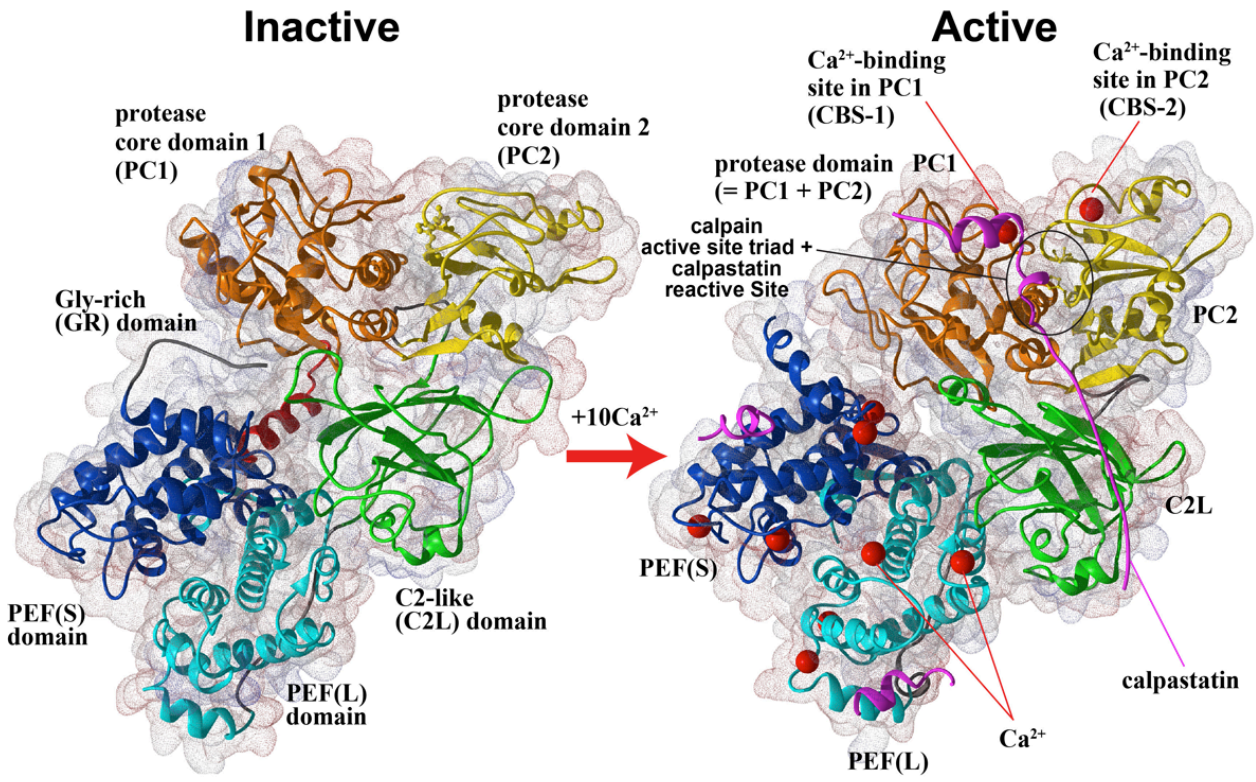


Figure 4. Schematic of the 3D structure of inactive human CAPN2 and Ca^{2+} activated rat CAPN2.

Unbound and calcium ion bound crystal structures of human and rat CAPN2 represented by cartoon 3D structures overlaid on 'surface' structures are shown. Binding of Ca^{2+} cause a conformational change resulting in activation of the catalytic triad. Image adapted from Sorimachi et al. 2011 (1).

Domain III shares a similar 3-dimensional structure to a C2 domain, but it shares little sequence similarity and is therefore referred to as a C2-like domain (C2-L). C2 domains are Ca^{2+} regulated lipid binding domains found in many enzymes. It is postulated that calpain translocation to phospholipid membranes is facilitated by the C2-L domain (25). However, there is also evidence that the N-terminal glycine rich hydrophobic domain V (GR) of the regulatory subunit might be of greater significance for membrane binding (25,26).

Domain IV is composed of a PEF domain. EF hands are helix-loop-helix motifs that typically bind calcium ions. However, the fifth EF hand in Domain IV binds with the fifth EF hand of another PEF domain in the regulatory subunit (domain VI), resulting in heterodimerisation of the large and small subunits (27,28). Atypical calpains do not have a PEF domain and therefore do not associate with a regulatory subunit (29).

In addition to the protease core domain atypical calpains also include a variety of repeated domains and other domains not found in typical calpains. A survey of calpain diversity in a representative sample of eukaryotic organisms revealed 24 different domains, including zinc fingers, microtubule interacting domains and glycine rich regions found in combination with the CysPc domain (8). The presence or absence of these domains are used to classify atypical calpains into different subfamilies (Figure 5): The PalB subfamily, which is found in humans, nematodes and other species, has two C2-L domains at its C-terminus (1). The SOH subfamily is characterised by presence of Zinc finger motifs at their N-terminus and a small optic lobe (SOL) homologue domain at their C-terminus. The DEK-1 subfamily has one C2-L domain and often an N-terminal trans-membrane domain, which might be subject to autolysis (9).

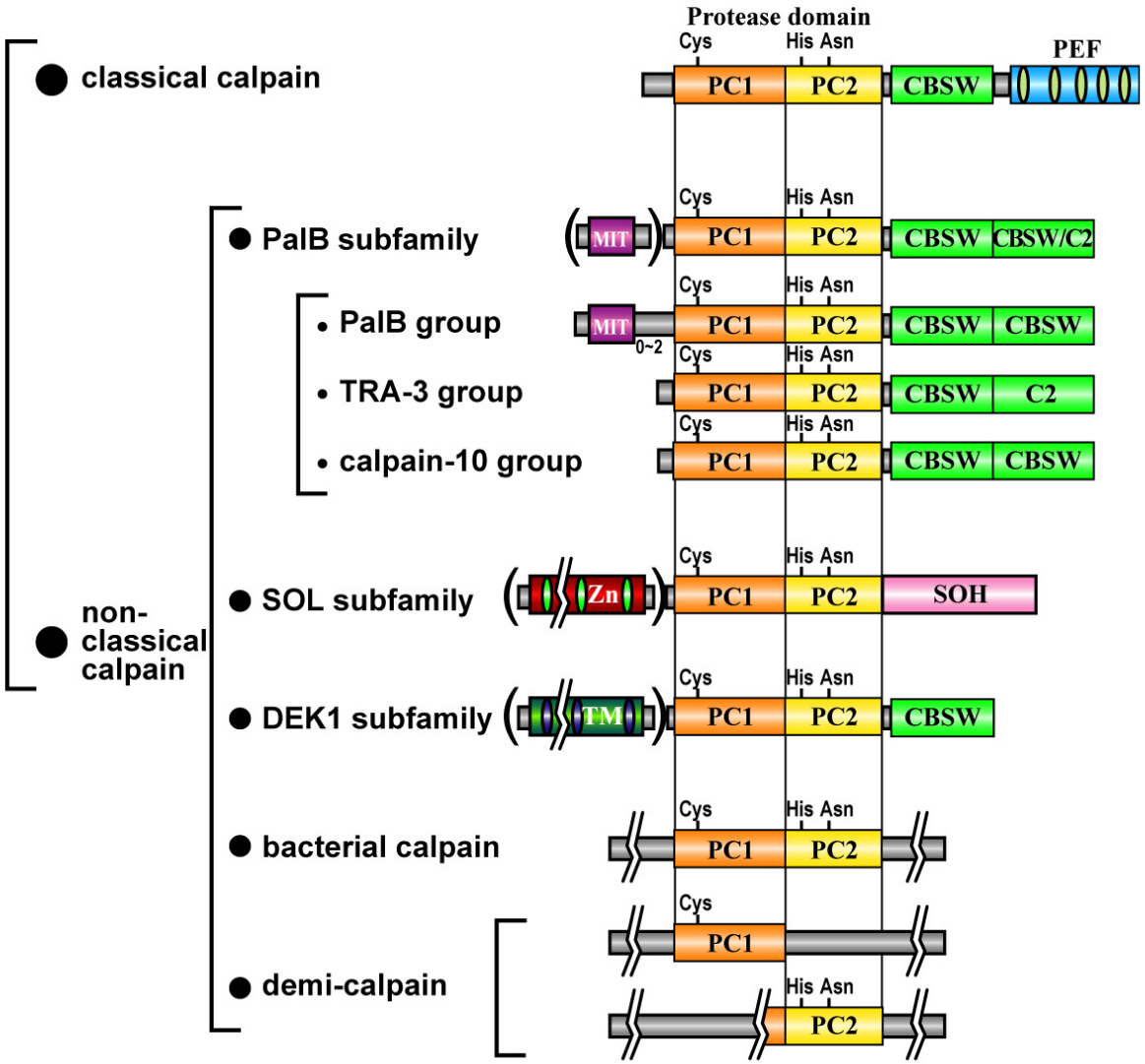


Figure 5. Classical or typical calpains and non-classical or atypical calpains and subfamilies.

Subfamilies are differentiated by the presence or absence of domains outside of the catalytic core PC1 and PC2 domains. Image adapted from Sorimachi et al. 2011 (30).

1.3 Substrate specificity of calpain.

Calpains have been proposed to act as 'modulator' proteases because they digest their substrates in a limited manner, affecting their longevity, localisation or activity. Calpain proteolysis can result in aberrant signal transduction or facilitate subsequent degradation (1). Specific examples of substrates modified by calpain proteolysis are discussed in section 1.4.

The substrate specificity of calpain is primarily determined by the 3-dimensional conformation of the substrate. Studies using peptide libraries to identify potential calpain substrates have identified consensus residues favoured by calpain proteases. However, the sequences identified using peptide libraries often differ from known substrate cleavage sequences (4). It is thus hypothesised that the substrate specificity of calpains is only weakly determined by primary sequence (4,17). For this reason, it has been difficult to identify calpain substrates and to predict sites of proteolysis. The deep cleft of the CysPC domain suggests that the substrate needs to be in an exposed loop conformation (4,24).

Over 100 substrates for calpains have been identified in vitro (10). These substrates can be grouped into four distinct types: cytoskeletal proteins, membrane-associated proteins, transcription factors, and kinases and phosphatases. Substrates identified in vitro are not necessarily in vivo physiological substrates. For example, fibronectin is degraded in vitro, but is unlikely to be an in vivo physiological substrate as it is an extracellular protein. Calpains are exclusively intracellular, except in pathological conditions (10).

1.4 Physiological roles of calpain

The conservation of calpain proteases across eukaryotic and some eubacterial phyla suggests that calpain function is important for eukaryotic life. The full-range of physiological functions of typical and atypical calpains is not well understood; however, It has been shown that calpains are involved in cytoskeletal remodelling, apoptosis, sex determination, signal transduction, gene expression regulation and the cell cycle (10,31). Most of this knowledge has been gained by using calpain inhibitors or genetic methods to reduce or enhance calpain activity. Calpain has been implicated to be a key mediator of p53. In a model of apoptotic cell death from DNA damage, it was observed that induction

of tumour suppressor p53 was reduced in cells that had been treated with MDL 28170, an active site targeting reversible membrane permeable peptidic calpain inhibitor. A similar reduction in p53 activity was obtained in cells derived from CAPNS (small regulatory subunit) knockout mice (32).

CAPN1 and CAPN2 were shown to have essential functions in CAPNS -/- knock out mice (33). Knockout of both CAPN1 and CAPN2 caused embryonic lethality after 11 days of gestation. Singly mutant CAPN1 -/- mice had no embryonic defects, but adults had an impaired platelet-platelet aggregation phenotype (34). These results suggest that CAPN1's function can mostly be replaced by CAPN2, and that CAPN1 has a specific signalling role in platelet aggregation. CAPN2 -/- mice failed to develop past the preimplantation stage (35), indicating that CAPN2 is physiologically more important than CAPN1. Histological samples of muscle tissue from CAPN3 -/- mice exhibited similar defects to those found in LGMD2a caused by mutations in the human CAPN3 gene (36).

CAPN1 and CAPN2 are inhibited by the only known endogenous calpain inhibitor calpastatin (Figure 6) (1,16). Calpastatin contains four inhibitory subunits that each inhibit calpain by binding the PEF domains IV and VI and CysPC domain. Calpastatin does not bind atypical calpains, which lack PEF-hand domains (37). NMR studies show that the calpastatin peptide normally assumes a random coil conformation (38). Calpastatin exists in several isoforms, although the significance of these isoforms is not understood (10).

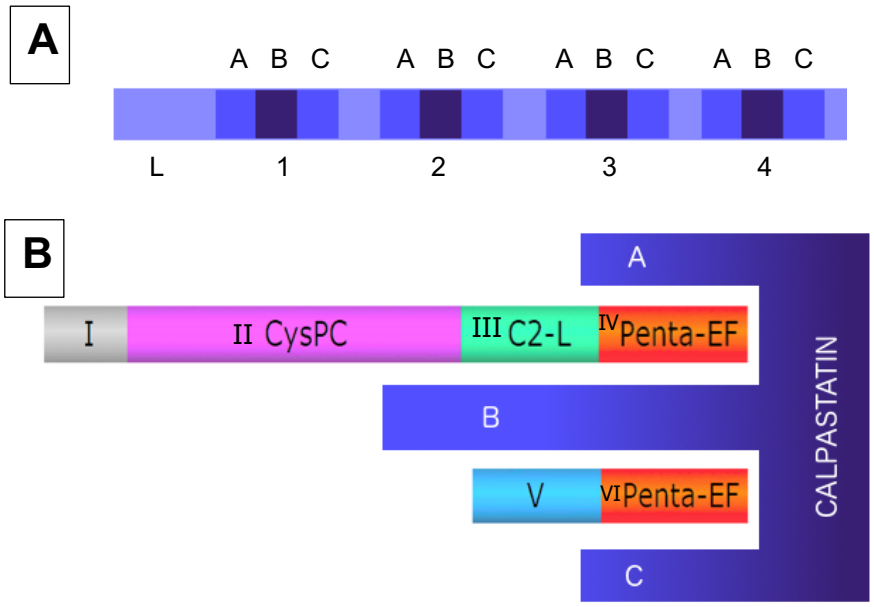


Figure 6. Calpastatin structure, Inhibition of typical calpains by calpastatin.
A: Calpastatin contains an L domain of unknown function and 4 repetitive inhibitory domains with A, B and C regions that can competitively bind and inhibit typical calpains (37). B: Calpastatin binds the penta EF-hand domains of typical calpains. Figure adapted from Chakraborti et. al 2012 (39).

Calpain inhibitors are classified into reversible and irreversible peptidic and reversible non-peptidic inhibitors. Many of these inhibitors are not specific to calpains and are capable of binding other cysteine and even serine proteases (10,40). Peptidic inhibitors include reversible inhibitors leupeptin and calpeptin, and irreversible inhibitor E-64, however these inhibitors are not selective for calpains. The non-peptidic small molecule inhibitor PD150606 inhibits calpain by an unknown mechanism by binding in a region outside of the active site (41). With the promise of therapeutic use, the identification of calpain inhibitors is an active area of drug discovery, with the aim of identifying calpain specific inhibition (42).

Cell motility is one of the most extensively studied processes involving calpains (31). Inhibition of calpain proteolysis has been shown to reduce focal adhesion complex

turnover, thus inhibiting cell motility and cell spreading. Focal adhesions are large protein complexes connecting the cytoskeleton with the extra cellular matrix (ECM) (31). CAPN1 and CAPN2 cleave components of adhesion complexes found at the rear of polarised cells, including talin and focal adhesion kinase (FAK). This results in detachment of the rear of the cell from the ECM helping facilitate cell movement (43,44).

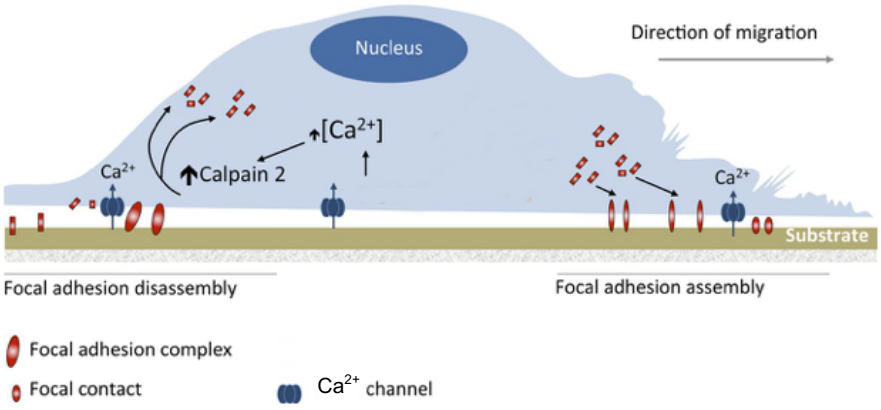


Figure 7. CAPN2 cleaves focal adhesion complex components facilitating cell migration.

Influx of calcium from calcium ion channels near to focal adhesion complexes activates CAPN2 which cleaves focal adhesion components including talin and FAK, allowing retraction of the rear of the cell from the extracellular matrix (44,45). Figure adapted from Saraiva et al. 2013 (45).

Calpains have an important role in cell-cycle regulation, and the crosstalk between autophagy and apoptosis (46). Autophagic cell death can occur as a fail-safe mechanism when apoptosis fails due to insufficient activation of caspase proteases (47). In human cells, overexpression of the autophagy factor ATG-5 increased autophagy and sensitised cells to undergo receptor and stress mediated cellular death (46). ATG-5 was shown to be cleaved by calpain to produce a functional product that translocates to mitochondria and inhibits the anti-apoptotic molecule Bcl-xL, triggering cytochrome c release and apoptotic cell death.

Calpains have been shown to inhibit autophagy; key autophagy proteins ATG5 and BEC-1 are cleaved by calpains (46,48). However, paradoxically calpain is also required for autophagy (49), showing that calpains interaction with cellular processes can be complex.

1.5 Calpain pathologies

1.5.1 Pathologies associated with typical calpains

The calpain-cathepsin hypothesis proposes that calpains promote necrotic cell death by permeabilising the lysosomal membrane resulting in cytosolic acidosis and release of cathepsin proteases (Figure 8). This hypothesis was proposed by Yamashima after it was found calpains congregate to lysosomal membranes after ischemic events (50). Support for this hypothesis was obtained in *C. elegans* after it was found that CLP-1 and TRA-3 were required for necrotic cell death in synergy with lysosomal (cathepsin) aspartyl proteases *asp-3* and *asp-4* (51).

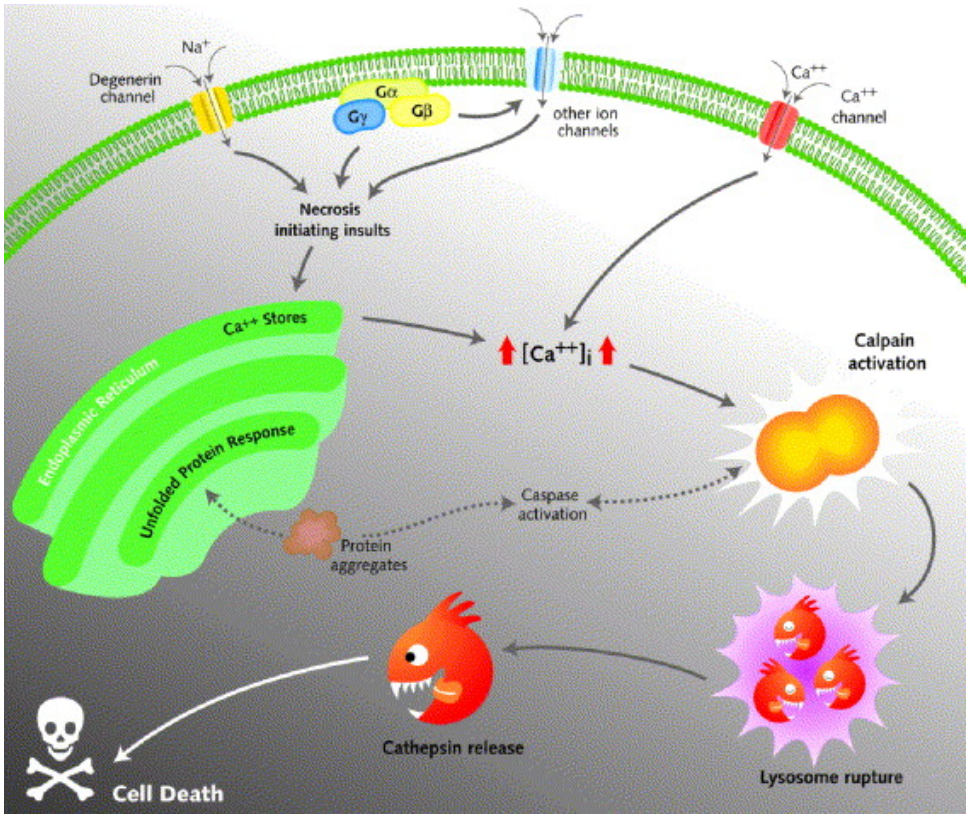


Figure 8. The calpain-cathepsin hypothesis.

Increase of cytosolic Ca^{2+} levels resulting from an influx from ion channels or intracellular stores after a necrosis triggering event, causes calpain activation. In turn calpains migrate to the lysosome where they permeabilise the lysosome and release cathepsin proteases. Cathepsins general degradation of cellular proteins results in necrotic cell death (50). Figure adapted from Syntichaki and Tavernarakis 2002 (51).

Diseases associated with calpains are classified into two types: genetic diseases associated with reduced or loss of calpain activity and diseases associated with over activity of calpains caused by elevated intracellular calcium levels or mutations that result in constitutive activity usually leading to degenerative pathologies.

Muscular dystrophies are denoted by the atrophy of muscle cells and tissues, and resulting progressive muscle weakness and degeneration (52). Nearly all muscular dystrophies involve the loss of Ca^{2+} homeostasis (29). The majority of muscular dystrophy diseases are caused by mutations in genes that encode structural proteins. However, LGMD2a is a form of muscular dystrophy that primarily affects the shoulder and pelvic musculature, and is caused by mutations in the gene encoding the skeletal muscle specific CAPN3 (11). Mutations that affect the ability of CAPN3 to degrade a protein substrate are hypothesised to be responsible for this disease (53). Increased apoptosis and changes in the localisation profile of nuclear factor κB (NF-κB) have been observed in LGMD2a patients (54). NF-κB is a transcription factor that drives expression of several survival genes (55). CAPN3 overexpression has also been identified as an aggravating factor in a mouse model of muscular dystrophy, showing the importance of CAPN3 regulation in normal physiology (54).

Human typical calpains CAPN8 and CAPN9 form a heterodimer in vivo. Both proteins are expressed predominantly in surface mucus cells of the stomach and intestine (56). Pit cells are essential in maintaining the stomach lining by secreting bicarbonate and neutral mucous. *CAPN8* and *CAPN9* knockout mice showed a significant increase in susceptibility to gastric mucosal injury. Expression of the *CAPN9* gene is down-regulated in gastric cancer tissue and thus CAPN9 is suspected to act as a tumour suppressor (56).

The majority of degenerative pathologies associated with calpains are those in which calpains are overactive due to disrupted Ca^{2+} homeostasis, resulting in elevated intracellular Ca^{2+} levels. Aging is one of the processes linked to increased Ca^{2+} disruption; however, many conditions can cause Ca^{2+} disruption including heat stress. Dysregulation of calpain activity is associated with several diseases including Alzheimer's disease, cataract formation, Duchenne and Becker's muscular dystrophy, myocardial infarction, multiple sclerosis and neuronal ischemia (10). Calpains have been identified as acting upstream of two key biochemical features of Alzheimer's disease; amyloid beta protein formation (senile plaques) and tau protein hyperphosphorylation resulting neurofibrillary tangles (57). Typical calpains have also been shown to cleave tau proteins and cleave hyperphosphorylated tau proteins with diminished efficiency (58). CAPN2 is hypothesised to act upstream of both of these pathological events, regulating the activity of neuronal developmental protein CDK5 kinase and GSKB kinase (59), and is therefore viewed as a potential therapeutic candidate for Alzheimer's disease (57,60).

CAPN2 cleaves alpha and beta crystalline following increase of calcium ion concentration of 20 fold caused by injury or aging (61), disrupting crystalline lattices that increase refractive index, resulting in cataract formation (62).

1.5.2 Pathologies associated with atypical calpains

Mutations in the atypical calpain CAPN10 gene are associated with an increased risk of having Type 2 diabetes mellitus (63). The disease is characterised by high blood pressure and high glucose levels in the blood caused by insulin resistance or inadequate insulin secretion. CAPN10 inhibition results in the decrease of insulin-stimulated uptake and also the decrease of GLUT4 translocation through cytoskeletal regulation (63).

The majority of diseases associated with calpains are attributed to aberration of typical calpains. Atypical calpains are less well characterised and may also be contributing to disease phenotypes. In this study I focus on pathologies associated with atypical calpains using *C. elegans* which expresses only atypical calpains.

1.6 *C. elegans* as a model organism

C. elegans is a free-living nematode worm adopted as a model organism by Sydney Brenner in an extensive forward genetic mapping study (64). Brenner chose *C. elegans* as a model to study the eukaryotic nervous system due to its relatively simple nervous system (302 neurons), transparency, and other advantageous characteristics (65) including: small size of adults (1 mm), inexpensive growth on *E. coli* (64), a rapid life cycle of 3.5 days at 20 °C, and hermaphrodite and male sexes that are convenient for genetic analysis.

Since Brenner initiated studies *with C. elegans*, several developments have increased its appeal as a model organism for biomedical research. The lineage and anatomy of the *C. elegans* hermaphrodite have been fully described, including the developmental fate of all 959 somatic cells found in hermaphrodite adults and 1031 somatic cells in adult males (65,66). *C. elegans* was also the first multicellular eukaryotic organism to have its whole genome sequenced in 1998, enabling bioinformatics methods for gene identification and reverse genetic studies of all *C. elegans* genes (67). More recently RNA interference and whole genome feeding libraries methods have been developed (68). As protein mass spectrometry technology has advanced, biochemical techniques for protein analysis have also been developed facilitating proteomic investigation (69). Due to the ease of generating transgenic lines, forward genetics, RNAi and high throughput chemical screens as well as the conservation in cell physiology, with 38% of *C. elegans* proteins having human homologues (70), *C. elegans* has been used as a disease model for neurodegenerative diseases such as Alzheimer's disease and Parkinson's disease as well

as muscular diseases and metabolic diseases (71). Several strains expressing human amyloid-beta protein in different tissues have been generated (71,72).

Whole genome sequence analysis of *C. elegans* has identified 10 genes encoding calpain proteins (5). The *C. elegans* atypical calpains CLP-1 and CLP-4 are expressed in body wall muscle cells (5). In this project, I have focused on identifying potential substrates of the *C. elegans* calpain-1 (CLP-1) protein. In addition, I initiated studies to understand the basis of the intracellular localisation of CLP-1 in body wall muscle. CLP-1 has a glycine rich N-terminus, an α-helix domain 1, a CysPC domain and a C2-like domain (Figure 2) (5). The similarities between *C. elegans* and human muscle cells are discussed below (73).

1.7 *C. elegans* muscle

There are several muscle groups found in *C. elegans* with distinct functions; body wall muscle used for locomotion, pharyngeal muscle for feeding, intestinal muscles for digestion, vulval muscle cells for egg-laying (74).

The fundamental contractile unit in *C. elegans* muscle is the sarcomere (Figure 9). Sarcomeres form repeated units in body wall muscle giving rise to the observed striated pattern, or form single units in non-striated muscles, such as the pharyngeal pumping muscle cells. Sarcomeres consist of actin thin filaments that are anchored to the cell membrane via attachments to dense bodies (homologous to vertebrate z-lines) (Figure 9). Actin thin filaments are interspersed with myosin thick filaments, which are anchored to the cell membrane by M-lines. M-lines form rows that are interdigitated with dense bodies. Attachment plaques are found at muscle-muscle junctions (74,75). There are a variety of different proteins found in the integrin attachment complexes (M-lines, dense bodies and attachment plaques) (Figure 9). *C. elegans* striated muscle has been used as a model for studying human muscle as they share close protein and structural homology (73).

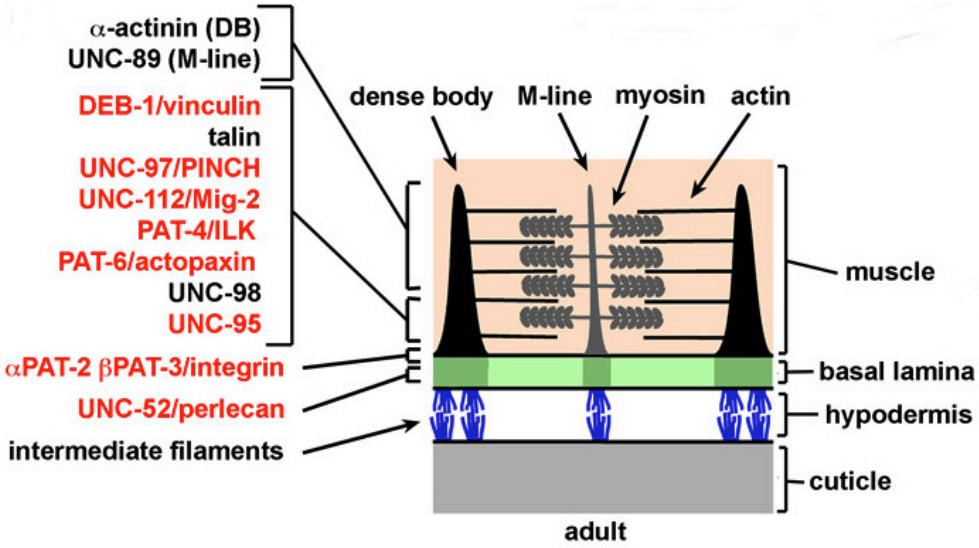


Figure 9. *C. elegans* Sarcomere.

Adapted from Moerman and Williams 2006 (76). The sarcomere comprises of dense body and m-line attachment complexes, which contain a variety of proteins some of which are important for normal muscle physiology (76). RNAi knockdown of sarcomeric proteins *unc-97* and *unc-112* induced mass protein degradation, myofibrillar and mitochondrial abnormalities and movement defects (77).

1.8 The atypical calpains of *C. elegans*

Name	Expression*	Notes
clp-1	Muscle, neuron, intestine, hypodermis	Required for neuronal necrotic cell death (51). Overexpression causes paralysis phenotype (5).
clp-2	Intestine	-
clp-3	-	-
clp-4	Muscle, neuron, hypodermis, excretory cell	-
clp-5/tra-3	Neuron, intestine, hypodermis, excretory cell	Required for correct sex determination (78). Required for neuronal necrotic cell death (51).
clp-6	-	-
clp-7	Neuron, intestine, hypodermis, excretory cell, seam cell	-
clp-8	-	-
clp-9	Pharynx, hypodermis, nervous system, tail, reproductive system	-
clp-10	-	clp-10 (ok2713) mutant has 50% reduction brood size (5).
clpr-1	-	Missing catalytic cysteine (5).
clpr-2	-	Missing catalytic cysteine (5).
clpr-3	-	Missing catalytic cysteine (5).
clpr-4	-	Missing catalytic cysteine (5).

Table 1. The calpains of *C. elegans*.
10 genes encoding atypical calpains are present in the *C. elegans* genome along with 4 *clp*-related genes that lack a complete catalytic triad and so do not have proteolytic activity (5). *Expression refers to evidence antibody staining of tagged reporters or fusion proteins (5,79).

Whole genome sequence analysis of *C. elegans* identified 10 genes encoding calpain proteins and 4 genes encoding calpain-like proteins, which have an incomplete catalytic triad (Table 1), all of which are lacking PEF domains.

TRA-3 (Calpain-5, CLP-5) is a *C. elegans* atypical calpain that plays a role in nematode sex-determination (78). TRA-3 proteolyses the membrane protein TRA-2A to produce a peptide that inhibits the masculinising protein FEM-3. Absence of TRA-3 causes the transformation of hermaphrodites into pseudomales. In this case, TRA-3 is seen as an important potentiator of its physiological substrate TRA-2A, which has feminising activity (78).

The *C. elegans* atypical calpains are expressed in a variety of tissues including neurons, intestine and body wall muscle (5). A systematic study of the calpains of *C. elegans* was performed by Joyce et al. (5). RNAi knockdown and deletion mutants of calpain genes yielded no observable phenotypes, suggesting either that there is functional redundancy between the *C. elegans* atypical calpains or that the functionality is difficult to discover through standard phenotypic measurements. *clp-1* and *clp-4* are the only genes expressed in body wall muscle and RNAi knockdown of *clp-1* or *clp-4* did not cause myofibrillar disorganisation based on phalloidin staining of actin fibres (5).

The atypical *C. elegans* calpain CLP-1 was overexpressed using the constitutively active muscle specific promoter of *unc-54* in the body wall muscle of transgenic animals to simulate the over activity of calpain (5). Animals overexpressing CLP-1 had an observable muscle phenotype; 1.5% of adults developed paralysis. Paralysis

was attributed to the proteolytic activity of CLP-1 because this phenotype was not observed in animals overexpressing a catalytically inactive version of CLP-1 with an alanine residue substituted for the catalytic cysteine residue. EGL-19 is a *C. elegans* calcium ion channel subunit protein expressed in body wall muscle cells important for normal muscle contractions (80). *egl-19(gf)* mutants have inappropriately elevated Ca^{2+} levels (80). When CLP-1 was overexpressed in *egl-19(gf)* mutants, 20% of adults displayed paralysis. Importantly, paralysis was suppressed when worms were treated with a Ca^{2+} ion channel inhibitor nemadipine-A, which specifically suppresses the activity of EGL-19 (5).

In a *C. elegans* model of neurodegeneration, RNAi reduction of the calpain proteases *clp-1* and *tra-3* or the aspartyl proteases *asp-3* and *asp-4* resulted in a reduction in the number of cells undergoing necrotic cell death. RNAi knockdown of both aspartyl and calpain proteases did not further reduce necrotic cell death, indicating that the aspartyl and calpain proteases acted in the same necrotic cell death pathway (51).

1.9 This project

Knowledge of calpains is based primarily on studies of typical calpains. The primary method of understanding the physiological role of a calpain is to identify substrates that are cleaved in vivo. This project focused on identifying atypical calpains substrates in body wall muscle using the model organism *C. elegans*. CLP-1 is required for necrotic cell death and overexpression has been shown to cause paralysis in adult *C. elegans* in a model of Duchenne muscular dystrophy (5,51).

My project had 3 aims:

1. To identify potential CLP-1 substrates in body wall muscle when CLP-1 is overexpressed in a *C. elegans* model of Duchenne muscular dystrophy.
2. To understand how CLP-1 is localised to sarcomere substructures in the body-wall muscle of *C. elegans*.
3. To determine if calpains are involved in heat stress.

2 Materials and Methods

2.1 Materials

Unless otherwise stated materials were obtained from Sigma-Aldrich Corporation.

2.1.1 Bacterial strains

Escherichia coli (*E. coli*) XL-1 blue cells: *recA1 endA1 gyrA46 thi-1 hsdR17 supE44 relA1 lac [F´ proAB lacIq ZΔM15 Tn10(Tetr)]*
E. coli OP50: Uracil auxotroph.
E. coli NA22: prototroph.

2.1.2 *C. elegans* strains

Strain	Genotype	Reference
Bristol N2	Wild type	(64)
PK1009	*dpy-5(e61) I;bli-2(e768) II;unc-32(e189) III*	(81)
PK1003	*unc-5(e53) IV; dpy-11(e224) V; lon-2(e678) X*	(81)
PK2969	*qyIs43 [pat-3::GFP + ina-1(genomic) + unc-119(+)]; crEX467.*	(82)
PK3085	*unc-54::clp-1::mRFP crEX468*	This study.
PK3113	*unc-54::mRFP crIs17 II*	This study.
PK3114	*unc-54::clp-1(C371A)::mRFP crIs18 X*	This study.
PK3118	*clp-1p::clp-1::GFP + pCoel::mRFP crEX484*	This study.
PK3142	*unc-119(ed3)*; qyIs43 [*pat-3::GFP* + *ina-1*(genomic) + *unc-119(+)*]; *unc-54::clp-1(C371A)::mRFP crIs17 II*	This study.
PK3143	*unc-54::clp-1(C371A)::mRFP*	This study.
PK3144	*unc-54::mRFP crEX491*	This study.
PK3146	*unc-54::clp-1Δ3::mRFP crEx499*	This study.
PK3149	*unc-119(ed3); qyIs43 [pat-3::GFP + ina-1(genomic) + unc-119(+)]; unc-54::clp-1::mRFP crIs18 X*	This study.
PK3150	*unc-119(ed3); qyIs43 [pat-3::GFP + ina-1(genomic) + unc-119(+)]; unc-54::mRFP crIs19*	This study.
PK2431	*clp-1(tm690)*	(5)
PK3163	*unc-54::clp-1::mRFP crIs19*	This study.
PK3228	*crEx501 [unc-54::clp-1Δ1::mRFP]*	This study.
BC10095	*sEx10095 [rCesF11C3.3::GFP + pCeh361]*	(83)

2.1.3 Plasmids

Name	Description	Reference
pPJ107	pPD30.38 vector with *unc-54::clp-1::myc* insert.	(5)
pPJ108	pPD30.38 vector with *unc-54::clp-1(C371A)::myc* insert.	(5)
pMC7	pBSKSII+ vector with *unc-54::clp-1::mRFP* insert.	(84)
pCFJ151	A Multiple Cloning site vector that is used for MosSCI insertion at ttTi5605 location. Contains homology arms consisting of genomic DNA flanking the insertion site.	(85)
pRF4	Encodes a mutant collagen *rol-6(su-1006)* that causes worms to have a dominant roller phenotype. Used as a marker for microinjection.	(86)
pLN1	pCFJ151 vector with *unc-54::clp-1(C371A)::mRFP* insert.	This study.
pLN3	pBSKSII+ vector with *unc-54::clp-1Δ3::mRFP* insert. Δ3 refers to a disruption mutation of domain 3 of CLP-1.	This study.
pLN4	pBSKSII+ vector with *unc-54::clp-1Δ1::mRFP* insert. Δ1 refers to a disruption mutation of domain 1 of CLP-1.	This study.
pLN5	pBSKSII+ vector with *unc-54::mRFP* insert.	This study.
pLN6	pBSKSII+ vector with *unc-54::clp-1Δ12::mRFP* insert. Δ12 refers to a disruption mutation of domains 1 and 2 of CLP-1.	This study.
pLN7	pBAD_myc_his_C with *clp-1* isoform a cDNA insert	This study.
pLN8	pBSKSII+ vector with *unc-54::clp-1Δ23::mRFP* insert. Δ23 refers to a disruption mutation of domains 2 and 3 of CLP-1.	This study.
pLN9	pBSKSII+ vector with *unc-54::clp-1Δ2::mRFP* insert. Δ2 refers to a disruption mutation of domain 2 of CLP-1.	This study.
pLN10	pBSKSII+ vector with *unc-54::clp-1Δ13::mRFP* insert. Δ13 refers to a disruption mutation of domains 1 and 3 of CLP-1.	This study.
pLN11	*clp-4* del mutant gRNA_1	This study.
pLN12	*clp-4* del mutant gRNA_2	This study.

2.1.4 Buffers and solutions

1X H100 buffer: 100 mM KCl, 1 mM $MgCl_2$, 50 mM (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) HEPES, 1 mM ethylene glycol tetraacetic acid (EGTA), 10% (w/v) glycerol. Prior to mixing with worms before homogenisation step add 1X Protease inhibitors (Roche), 1 mM phenylmethylsulfonyl fluoride, 0.5 mM DTT and 0.05% (v/v) NP-40.

10x TBE: 0.9 M TRIS-base, 0.9 M boric acid, 0.02 M EDTA.

1.33x Gibson master mix: 1.33x isothermal buffer, Taq DNA ligase (New England Biolabs) 1000 units, T5 exonuclease (New England Biolabs) 20 units, Phusion DNA polymerase (New England Biolabs) 12.5 units. For 375 µl of Gibson master mix: 50 µl taq ligase (40U/µl), 100 µl 5x isothermal buffer, 2 µl T5 exonuclease (1U/µl), 6.25 µl phusion polymerase (2U/µl), nuclease free water 216.75 µl.

1.5X H100 buffer: 150 mM KCl, 1.5 mM $MgCl_2$, 75 mM HEPES, 1.5 mM EGTA, 15% glycerol, 2X Protease inhibitors (Roche).

2x SDS sample buffer: 0.125 M TRIS-HCL pH 6.8, 20% (w/v) glycerol, 4% SDS, 3 mM bromophenol blue, 0.2 M β-mercaptoethanol added prior to use.

2x Nematode freezing solution: 0.1 M NaCl, 0.05 M KH2PO4, 30% (w/v) glycerol, 5.6 mM Sodium hydroxide (NaOH). Autoclaved then 0.3 mM MgSO4 added.

5x isothermal buffer: 25% (w/v) polyethylene glycol-800, 0.5 M Tris-hydrogen chloride pH 7.5 1.5ml, 50 mM magnesium chloride ($MgCl_2$), 50 mM dithiothreitol (DTT), 1 mM deoxyadenosine triphosphate (dATP), 1 mM deoxythymidine triphosphate (dTTP), 1 mM deoxycytidine triphosphate (dCTP), 1 mM deoxyguanosine triphosphate (dGTP), 5 mM nicotinamide adenine dinucleotide (NAD).

50x TAE: 2 M tris(hydroxymethyl)aminomethane) (TRIS-base), 1.9 M glacial acetic acid, 0.05 M ethylenediaminetetraacetic acid (EDTA) pH 8.

Blocking buffer: 5% (w/v) non-fat dried skimmed milk (Marvel), 0.1% Tween 20 (Melford), in PBS.

Phosphate Buffered Saline (PBS): 0.14 M NaCl, 2.5 mM Potassium chloride KCl, 8 mM Sodium dihydrogen phosphate, 16 mM KH_2PO_4.

PBS-T: PBS, 0.1% (v/v) Tween-20.

Ponceau stain: 0.1% (w/v) Ponceau S (P3504), 0.83 M acetic acid.

Single worm PCR Lysis buffer: 50 mM Potassium chloride (KCl), 10 mM Tris-HCl (pH 8.3), 2.5 mM Magnesium Chloride ($MgCl_2$), 0.45 % Nonidet P-40, 0.45% Tween 20, 0.01% Gelatin, 0.01% Proteinase k.

Sodium dodecyl sulphate (SDS) running buffer: 0.025 M TRIS-base, 0.19 M glycine, 0.1% (w/v) SDS.

TRIS-base/EDTA (TE) buffer: 0.01 M TRIS-Hydrochloric acid (TRIS-HCl) pH 8, 0.1 mM EDTA.

Western Transfer Buffer: 25 mM TRIS-base, 0.19 M Glycine, 20% (v/v) methanol.

2.1.5 Media

Premixed Luria-Bertani medium (LB) was prepared as directed (Sigma, L3022).

1.5% LB agar plates were prepared by adding agar (Melford Chemicals, M1002) to LB medium prior to autoclaving.

$CaCl_2$ competent cells solution: 1 M $CaCl_2$, 1 M piperazine-N,N'-bis(2-ethanesulfonic acid) (PIPES), 15% v/v glycerol.

2.1.7 Antibiotics

When required, antibiotics were added directly to media in the following final concentrations:

Ampicillin: 100 µg/ml.

Tetracycline: 12 µg/ml.

2.1.8 Antibodies

Nanotrap: RFP-Trap_A (ChromoTek, cat. no: rta-20)

2.1.8.1 Primary antibodies

polyclonal rabbit anti-RFP antibody (MBL, cat. no: PM005). Dilution: 1:1000

MH4 mouse anti-intermediate filament subunit antibody (DSHB). Dilution 1:200

MH24 mouse anti-vinculin Dilution 1:200

MH25 mouse anti-beta integrin Dilution 1:200

Mouse monoclonal anti-tubulin antibody (Sigma, cat. no: T5168) 1:1000

2.1.8.2 Secondary antibodies

ECL goat anti-rabbit IgG, HRP-linked whole Ab. Dilution: 1:5000 (Thermo)

ECL goat anti-mouse IgG, HRP-linked whole Ab. Dilution: 1:5000 (Thermo)

2.2 Molecular Biology Methods

2.2.1 Agarose gel electrophoresis (Diagnostic and Extraction)

1% (w/v) Agarose (Invitrogen) solutions were made with either 0.5x TBE or 1x TAE buffer, heated in a microwave and stored at 65°C ready for pouring.

TBE gels were used for diagnostic purposes. Samples were mixed with 6x ficoll dye and electrophoresed at 65 V for 25-35 min in 0.5x TBE buffer alongside a DNA molecular weight ladder (Promega G5711) (87). The gel was then stained in 0.5 µg/ml ethidium bromide (Bio-Rad) for >20 min before being visualised in a Gene Genius gene-doc system using GeneSys software.

TAE gels were used to purify DNA fragments; samples were mixed with 4 µl of ficoll-lite before being electrophoresed at 55 V for 1 hour in 1x TAE buffer (87). The gel was then stained with 1x SYBR Green (Thermo Fisher Scientific) for >30 min. DNA bands were visualised using a dark reader (Clare chemical) and the gel containing the band was excised using a razor blade. DNA was then extracted from the gel slice using a column-based QIAquick Gel Extraction Kit (Qiagen, cat. no: 28706) according to manufacturer's instructions.

2.2.2 Alkaline Lysis miniprep

Single colonies were inoculated in 2 ml LB broth containing 100 µg/µl ampicillin and placed in an Innova shaking incubator (New Brunswick) at 37°C for 16-18 hours. 1.5 ml of each culture was harvested, supernatant discarded, and pellet resuspended in 100 µl of GTE and vortexed. 100 µl of 1% sodium dodecyl sulphate (BDH laboratory

supplies) / 0.2 M sodium hydroxide (AnalaR) was added and mixed by shaking. 150 µl of 3 M potassium acetate / 2 M hydrogen acetate pH 4.8 was then added and mixed again by shaking. The mixture was spun at 15,000 xg for 3 min and supernatant transferred to a fresh Eppendorf tube containing 250 µl of isopropanol before being spun for another 3 min at 15,000 xg. The pellet was then washed with 100 µl of -20°C 70% ethanol before being resuspended in 50 µl of TE buffer. Adapted from (88).

2.2.3 Bacterial transformation of plasmids

E. coli XL-1 blue competent cells were thawed by hand and immediately placed on ice. 0.1 – 0.5 µg/µl DNA was added to a pre-chilled glass tube with 100 µl competent cells, this mix was kept on ice for 20 minutes before being heat shocked for 2 minutes at 42°C, then cooled on ice for 2 minutes before adding 250 µl LB broth and incubating at 37°C for 45-60 minutes. Transformation mixture was spread on LB agar plates containing 100 µg/ml ampicillin and incubated at 37°C for 18-20 hours. For blue/white selection 60 µg/ml X-gal (Melford) dissolved in N,N-dimethylformamide and 0.12 mM Isopropyl β-D-1-thiogalactopyranoside (Melford) dissolved in Milli-Q H_20 was spread on the ampicillin plates prior to spreading of the transformed bacteria (89).

2.2.4 DNA restriction digest

DNA was digested using the appropriate restriction endonucleases (New England Biolabs) in solutions containing buffer and BSA for times and temperatures recommended by the manufacturer.

2.2.5 Ethanol precipitation

DNA was concentrated by adding 50 - 150 µg/ml glycogen (Promega), 0.1 volume of 3 M Sodium acetate (Analar), 2 volumes of 100% ethanol and cooled in a -20°C freezer for >2 hours. The mixture was then centrifuged at 15,000 x g for 3 min and the supernatant removed using a pipette. The pellet was washed in -20°C 70% ethanol before being resuspended in 10 µl of either 1X TE buffer, 0.5X TE buffer or Milli-Q H20 depending on subsequent use2.

2.2.6 Gibson isothermal Assembly

Primers fused to synthesise PCR products for Gibson isothermal assembly reactions were designed with the help of the online NEBuilder tool (90). 10-100 ng of purified PCR products containing homologous overlaps was added to make up 5 µl DNA mix, which was added to 15 µl of Gibson assembly 1.33x master mix to make a final volume of 20 µl. The reaction was assembled on ice, before being incubated at 50°C for 1 hour in a PTC-225 Thermocycler (MJ Research) (91).

2.2.7 Preparation of $CaCl_2$ competent cells

E. coli XL-1 blue was streaked onto a LB-agar plate containing 10 µg/ml tetracycline and incubated at 37°C overnight. Single colonies were used to inoculate vials containing 3 x 5 ml LB broth, which were then incubated at 37°C overnight in a shaking incubator. 5 ml of overnight culture was used to inoculate 420 ml of LB broth containing 10 µg/ml tetracycline in a 2 L flask, shaken at 200 rpm in an Innova shaking incubator (New Brunswick) at 37°C until OD_{690} reached 0.375.

The culture was then split into ice cold 50 ml Falcon tubes, cooled on ice for 10 min and centrifuged at 3200 x g at 4°C for 7 min. The supernatant was discarded, and

the pellet washed with 10 ml of ice cold $CaCl_2$ competent cells solution, this spin/wash step was then repeated, and the cells left to chill on ice for 30 min. The cells were then pooled into four 50 ml Falcon tubes, spun at 1250 x g at 4°C for 5 min, supernatant discarded and pellet resuspended in 4 ml of ice cold $CaCl_2$ competent cells solution. Competent cells were then snap frozen with liquid nitrogen and stored at -80°C. Method adapted from (92).

2.2.8 Polymerase chain reaction (PCR)

High fidelity PCR was carried out using Phusion (New England Biolabs) according to manufacturer's protocol. PCR reactions of 50 µl volume were prepared containing: 1x HF buffer (New England Biolabs), 0.5 µM forward primer, 0.5 µM reverse primer, 50-100 ng template plasmid DNA, 200 µM deoxyribonucleotide triphosphates (dNTPs), 1 unit Phusion polymerase (cat. No: M0530, New England Biolabs) and MilliQ water up to 50 µl. PCR was carried out using a PTC-225 thermocycler (MJ research) using the following conditions: 98°C for 30 seconds, 25-35 cycles of 98°C for 10 seconds, 50-65°C for 20 seconds, 72°C for 15-30 seconds per kilobase of template followed by a final extension period at 72°C for 10 minutes.

2.2.9 PureYield DNA miniprep

Single colonies were grown in 5 ml LB containing 100 µg/µl ampicillin in an Innova shaking incubator (New Brunswick) at 37°C for 16-18 hours. 4.5 ml of bacteria was harvested and DNA was extracted using the plasmid Miniprep System kit (PureYield™, cat. No: A1222, Promega) following manufacturer's instructions.

2.2.10 Site directed deletion mutagenesis

Primers were designed and used in a standard PCR to amplify the plasmid region flanking the targeted deletion site, thus excluding the deletion region from the amplicon (adapted from (93)). 2 units of DpnI (New England Biolabs) was added directly to the PCR product and incubated at 37°C for 30-60 min. 4 µl of the DpnI digested PCR product was used in a 20 µl ligation reaction to circularise the PCR product, containing; 1x T4 kinase buffer (New England Biolabs), 10 units T4 polynucleotide kinase (New England Biolabs), 1 unit T4 ligase (New England Biolabs) and MilliQ water up to 20 µl. The ligation mix was incubated at room temperature for 20 min and product used to transform competent cells.

2.3 Protein methods

2.3.1 Bradford Assay

Using the Pierce™ Coomassie (Bradford) Protein Assay Kit (cat. Number: 23200, Thermo Scientific) a BSA (diluted in relevant buffer) standard concentration against absorbance at 595 nm curve was made. Serial dilutions (0.5x, 0.25x, 0.125x) of sample were plotted against this curve to determine sample concentration as per manufacturer's instructions.

2.3.2 SDS-PAGE gel

10% Acrylamide sodium dodecyl sulphate polyacrylamide gels were prepared in a Mini-PROTEAN® II Electrophoresis Cell (Bio-Rad). The gel was covered in water saturated butanol and left to set. Once set, a 4.75% acrylamide stacking gel was prepared on top of the resolving gel and comb inserted providing wells for samples

to be added. Protein samples were prepared by mixing with 2X SDS loading buffer and heating at 95°C for 10 mins prior to loading onto gel. The samples were electrophoresed at 100 V in 1X SDS-PAGE running buffer until the bromophenol blue dye had ran out the bottom of the gel (93).

2.3.3 Silver Staining

Polyacrylamide gels were stained using Pierce™ Silver Stain Kit (cat. Number: 24612, Thermo scientific) according to manufacturer's instructions.

2.3.4 Western blotting and antibody staining

Protein samples run on SDS-PAGE gels were transferred to nitrocellulose membranes (Amersham Protran 0.45 NC, cat. No.: 10600002, GE Healthcare) using a Mini-Protean II Cell Gel System (Bio-Rad) containing 1 litre of Western transfer buffer following the method of Towbin (94). The transfer was run at constant 40 mA at 4°C for 16-18 hours. The membrane was Ponceau S stained to assess protein transfer, then rinsed with MilliQ water to remove stain. The nitrocellulose membrane was then washed in 20 ml of blocking buffer and rocked for 30 min. The blocking buffer was removed, and the membrane was washed in 5 ml of PBS-T for 3x 5 min. Primary antibody diluted in blocking buffer was incubated with the membrane in a falcon tube for 1 hour. Nitrocellulose membrane was washed in PBS-T for 3x 5 min and the membrane was then incubated for 1 hour with secondary antibody diluted in blocking buffer before being washed for 15 min in PBS-T and then a final 3x 5 min wash in PBS-T. Membranes were then processed using SuperSignal West Pico Chemiluminescent substrate (Thermo Scientific, cat. No.: 34087) according to

manufacturer's instructions before detection by Odyssey Fc Dual mode imaging system (Li-COR) with exposure of 2-10 minutes.

2.4 *C. elegans* Methods

2.4.1 CRISPR/cas9-mediated mutagenesis

gRNA cloning (pLN11 and pLN12)

3 µl of each 100 mM Complementary gRNA oligonucleotide was added to 10 µl 5X annealing buffer mixed with 44 µl of ddH2O to make a total volume of 50 µl Annealing mix was then heated to 95°C and allowed to cool to room temperature (~45 min).

1 µl of annealed primers was then added to a 5 µl total volume ligation mixture containing: 1 µl of 50 ng/µl Bsa1 cut pRB1017, 2.5 µl of 2X rapid ligase buffer (Promega), 0.5 µl T4 ligase (Promega). The ligation mixture was incubated for 2 hours at room temperature and transformed into 100 µl of XL-1 Blue cells.

Colonies from transformation were screened for recombinants using M13_F and the specific gRNA antisense primer by bacterial colony PCR. Recombinant colonies were then extracted via plasmid miniprep and sequenced using M13_F primer.

Microinjection

Young adult PK3161 [*clp-1(tm690)*] worms were injected with the following CRISPR mixture: pDD162 50 ng/µl, pJA58 25 ng/µl, PK1559 500 nM, pLN_11 30 ng/µl, pLN_12 30 ng/µl, PK1822 500 nM. After microinjection worms were placed on single OP50 seeded plates and grown at 20°C.

Screening

5-6 days post microinjection, plates containing injected worms and their progeny were scored for: survival of parent, progeny displaying roller phenotype (count), progeny displaying dumpy phenotype (count). Lines with high numbers of roller and dumpy worms were considered ‘jackpots’ and were chosen for further screening. Progeny from ’jackpot’ plates were singled (40 – 75) on OP50 seeded plates, grown for 1-2 days at 20°C and allowed to lay progeny and subsequently screened by PCR for presence of the deletion mutation using primers PK1816 and PK1817.

2.4.2 DNA Microinjection of worms

Transgenic animals expressing extrachromosomal arrays were generated using the microinjection method as described by Mello et. al (86) . microinjection mixes contained 10 µg/ml of experimental plasmid, a co-expression marker (either 80 µg/ml pRF4 or 40 µg/ml pCoel::mRFP) that acts as a marker for successful microinjection and N2 genomic DNA to a total concentration of 90-100 µg/µl. Once injected, worms were placed on separate *E. coli* OP50 seeded plates. F1 progeny expressing the marker phenotype (roller) were placed on separate *E. coli* OP50 seeded plates and F2 progeny expressing the marker phenotype from these plates were used to establish stably transmitting lines.

2.4.3 Freezing of worms for long term storage

For long term storage, *C. elegans* strains were kept at -80°C or in liquid nitrogen (64). To freeze a strain, 1.5 ml of M9 buffer was used to wash off each of three starved plates containing L1 worms into a 15 ml falcon tube, which was then spun at 1200 rpm for 2 min in a centrifuge. Worms were pelleted at the bottom of the falcon tube; the supernatant was removed before 1 ml of 2x nematode freezing solution and 1 ml of M9 buffer was added. Mix was gently swirled before it was aliquoted into cryotubes (Nunc) which were subsequently placed inside a room temperature Mr Frosty TM freezing container (Thermo Scientific), which was placed inside a -80°C freezer and left for >24 hours. Frozen worms were thawed for use by hand and then poured onto an *E. coli* OP50 seeded NGM plate and left to recover in a 15°C incubator.

2.4.4 General maintenance of *C. elegans*

All strains of *C. elegans* used were cultivated in the lab on bacterial lawns of *E. coli* OP50 on single peptone Nematode growth medium (NGM) plates (referred to as seeded plates) and maintained at 15-25°C, according to the methods of Brenner (64). Integrated transgenic and N2 strains were maintained by chunking sections of a starved plate onto new plates. Worms expressing extrachromosomal arrays typically have low genetic transmission, so these worms were maintained by picking young adults using platinum wire with the associated phenotype onto new plates.

2.4.5 Heat stress assay

To synchronise animals, 20-30 adult worms were picked to an *E. coli* OP50 seeded NGM plate and allowed to lay eggs for 3 hours before being picked off, the worms from this 'timed egg laying' plate were considered synchronous.

30 Synchronised young adult worms were picked to a pre-warmed plate and subjected to heat stress at 35°C for 2 hours. Immediately after the 2 hour period worms were returned to 20°C and screened for paralysis and death and any abnormalities. Worms were screened again 2 hours later and every 24 hours for 3 days (from t=0) (95). Significance testing between strains was performed using Fisher's exact test.

2.4.6 Integrating extra-chromosomal arrays using irradiation

50 L4 transgenic worms expressing the extrachromosomal array of interest were irradiated with 48 Gy from a ^{137}Cs source. Irradiated worms (P0) were allowed to recover for 24 hours before groups of 2-4 irradiated animals were placed on *E. coli* OP50 seeded plates. Once grown, 10-20 F1 young adult progeny from each of these plates were picked individually onto new seeded plates (approximately 200-250 F1 were picked in total). F1 progeny (F2) were allowed to develop before being scored based on percentage of F2 with transgenic phenotype. F2 from the F1 plates with the 10 highest transmission rates were then picked on to fresh seeded plates and strains with 100% transmission frequency were identified as integrated. As following the method described by Macmorris (96). Integrated strains were outcrossed with N2 worms five times in order to remove extraneous mutations.

2.4.7 Liquid Culture

5-8 uncontaminated 5 cm diameter NGM plates containing just starved *C. elegans* strain of interest were collected by washing plates with 5 ml of M9 buffer, allowing plates to stand for one minute and then transferring the wash into 15 ml falcon tubes. The worms suspended in M9 were centrifuged at 1200 rpm for 2 min at 4°C and supernatant discarded. The pelleted worms were washed twice more in M9. The worms were used to inoculate a 500 ml S-Basal complete medium containing *E. coli* NA22 in a 2 L volumetric flask. The inoculated media was then incubated in a shaking incubator at 20°C at 180-200 rpm for 3-4 days, until the majority of worms were young adults. The worms were harvested by settling the liquid culture in a 1 L glass cylinder partially embedded in ice for 4 hours. The media was then aspirated until there was approximately 50 ml left in the cylinder, the remaining liquid was transferred to 15 ml falcon tubes and the cylinder washed with M9 to elute remaining worms and the wash transferred to a 15 ml falcon tube. The suspended worms were centrifuged at 1200 rpm for 2 min at 4°C, supernatant discarded, worm pellets resuspended in ice cold M9, pooled together and centrifuged again at 1200 rpm for 2 min at 4°C to yield 0.5 – 5 ml of packed worms. Packed worms were distributed into 500 µl aliquots, mixed with an equal volume of 1X lysis buffer containing 2X protease inhibitor cocktail (Roche, Cat number: 04693159001) and then snap frozen individually in liquid nitrogen to form 1 ml aliquots of worm pellets which were stored at -80°C (97).

2.4.8 Mapping of transgenes to chromosomes

To determine which chromosome an integrated transgene was located on in a strain, the strain was crossed with mapping strains PK1003 [*dpy-5(e61) I;bli-2(e768) II;unc-*

32(e189) III] and PK1009 [*unc-5(e53) IV; dpy-11(e224) V; lon-2(e678) X*] respectively. PK1003 is homozygous for scorable phenotypes dumpy, blistered and uncoordinated. PK1009 is homozygous for scorable phenotypes uncoordinated, dumpy and long. 3-4 progeny F1 heterozygous worms were picked onto single plates. F2 worms were then picked to new plates respectively based on segregated plate phenotype, and the number of worms displaying the transgenic phenotype and without the transgenic phenotype were counted. Exclusion of the transgenic phenotype from F2 worms that are homozygous for a plate phenotype suggests that the transgene is present on that particular chromosome.

2.4.9 Microscopy

2.4.9.1 Fluorescence microscopy

Fluorescence from worms expressing GFP and RFP were assayed using a Leica MZ FLIII fluorescence stereomicroscope using appropriate filters.

2.4.9.2 Confocal microscopy

Worms of interest were sedated using 1 mM levamisole 20 minutes prior to being mounted on set 3% agarose on glass slides for viewing using a Leica SP5 confocal system (with resonant scanner and hybrid detectors) (98). Fluorescent, DIC and z-stack images were captured using Leica Application Suite Advanced Fluorescence (LAS AF) and analysed using Volocity 6.3 (Improvision, Perkin Elmer) and ImageJ.

2.4.10 Setting up worm crosses

9-12 L4 males of a transgenic non-Unc strain of interest were put on a *E. coli* OP50 seeded plate with 2-3 L4 hermaphrodites of desired cross strain and progeny were assessed for cross phenotype (99).

2.4.11 Single Worm PCR

Single worms were picked in to a PCR tube containing 2.5 µl of freshly prepared single worm PCR lysis buffer. The PCR tube was then placed in the -80°C freezer for >2 hours. After freezing, a drop of mineral oil was added to the PCR tube before it was incubated at 60°C for 1 hour and then at 95°C for 20 minutes in a PTC-225 thermocycler (MJ research). PCR GoTaq reagents; 5 µl 5x GoTaq buffer (Promega), 1.5 µl 25 mM $MgCl_2$, 0.5 µl forward primer 10 µM, 0.5 µl reverse primer 10 µM, 0.2 µl 25 mM Deoxyribonucleotide triphosphates (dNTPs), 0.125 units of GoTaq polymerase (Promega), 14.7 µl MilliQ H_20 were then added directly to this PCR tube and PCR carried out according to manufacturer's instructions.

2.4.12 Whole worm Immunoprecipitation

1 ml of snap frozen worms in H100 lysis buffer pellets were ground to a powder using a liquid nitrogen cooled pestle and mortar. The powder was thawed out at room temperature and immediately placed on ice and mixed with an equal volume of ice cold 1.5X lysis buffer containing 1x Protease inhibitor cocktail (Roche, Cat number: 04693159001), 2 mM PMSF, 1 mM DTT and 0.1% (v/v) NP-40. This mix was then homogenised 30 times on ice using an ice cold Dounce homogenizer. The

homogenised mix was then centrifuged at 13.2 Krpm at 4°C for 10 min, the supernatant transferred to a new microcentrifuge tube and centrifugation/supernatant extraction repeated twice more. 20 µl of RFP nanotrap bead slurry (ChromoTek) was equilibrated in 500 µl of ice cold H100 buffer, centrifuged at 2500x g for 2.5 min at 4°C and washed twice more in H100 buffer. Before incubation with the beads the lysate concentration was determined by Bradford assay and diluted to 2 mg/ml. 1 ml of HSS was added to the equilibrated beads and incubated at 4°C for 2 hours with constant mixing. After incubation period, incubation mix was centrifuged at 2500x g for 2 min at 4°C, 50 µl of supernatant was transferred to new Eppendorf and kept as unbound fraction and rest of supernatant discarded. Beads were resuspended in 500 µl of ice cold H100 buffer with 1x Protease inhibitors, centrifuged at 2500x g for 2.5 min at 4°C, resuspension and centrifugation steps were repeated twice more (100). Supernatant was discarded, and beads were then sent to a proteomics facility for mass spectrometry analysis or boiled in 2X SDS sample buffer for 10 min and pelleted and used in gel based analysis. For gel based analysis, immunoprecipitation fractions were loaded equally based on volume.

3 Results

3.1 Identifying potential CLP-1 substrates in body wall muscle by co-immunoprecipitation and mass spectrometry.

Overview

Overexpression of the atypical calpain CLP-1 in body wall muscle was previously shown to cause a paralysis phenotype in a percentage of adult worms carrying the CLP-1 transgene in a *C. elegans* model of Duchenne muscular dystrophy (5). To extend this observation I sought to identify the structural proteins in muscle that might be proteolyzed by CLP-1. I will describe how I attempted to find substrates of the atypical calpain CLP-1 by co-immunoprecipitation (co-IP) followed by liquid chromatography tandem-mass spectrometry (LC-MSMS) as summarised in Figure 10. The identification of potential substrates was based on differential co-IP using strains expressing either catalytically inactive CLP-1(C371A)::mRFP or mRFP alone in muscle.

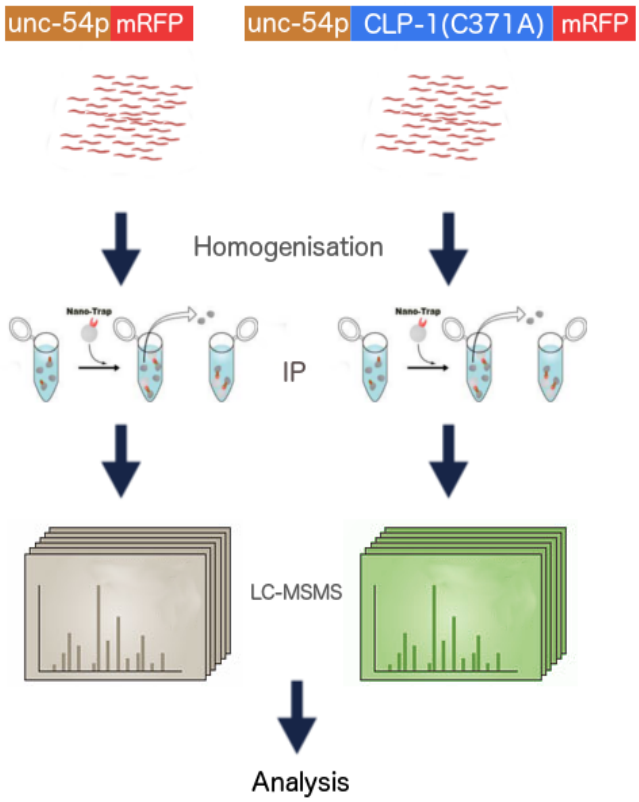


Figure 10. Work flow of proteomic investigation into CLP-1 substrates.

Transgenic animals were grown in batches then subsequently lysed. Immunoprecipitations of mRFP tagged proteins from lysate was performed, and co-immunoprecipitated proteins were then identified by LC-MSMS. Analysis filtered out CLP-1 specific binding partners by comparison to an mRFP control.

Alongside the proteomics experiment, it was also investigated if mammalian typical calpain substrates beta-integrin (PAT-3) (101), vinculin (DEB-1) (102) and desmin (IFA-1,2,3) (103) might be potential substrates of CLP-1 by co-immunoprecipitation and western blotting with antibodies specific to these proteins.

3.1.1 Construction of transgenic *C. elegans* strains expressing catalytically inactive CLP-1(C371A)::mRFP and mRFP alone

A catalytically inactive mutant of CLP-1 named CLP-1(C371A), which has a mutation causing substitution of the cysteine residue of the catalytic triad to an alanine, was used in this screen. CLP-1(C371A) was used to prevent cleavage of potential target substrates prior to co-IP and LC-MSMS. Previous work by Joyce et al. (2012) established that CLP-1::mRFP produced a similar phenotype to CLP-1::MYC when overexpressed in *C. elegans* body wall muscle (5), indicating that a C-terminal mRFP tagged CLP-1 retains its proteolytic activity. Therefore, it was concluded that C-terminal mRFP CLP-1 fusion protein would retain its ability to bind CLP-1 substrates, and these substrates would be co-immunoprecipitated with CLP-1::mRFP. A mRFP affinity tag was selected to tag CLP-1. mRFP fusion proteins are suitable targets for immunoprecipitation using RFP nanotrap (ChromoTek) (100). An mRFP tag was also chosen as the intracellular localisation pattern of CLP-1::mRFP could be analysed by in situ fluorescence microscopy.

The plasmid clone pLN1, which expresses CLP-1(C371A)::mRFP under the control of the muscle constitutive promoter of *unc-54* (104) (Figure 11), was generated by Gibson assembly using four PCR fragments from plasmids PCFJ151, pMC7 and pPJ108 (Figure 12). PCFJ151 is a worm expression vector (85), pMC7 contains wildtype cDNA of *clp-1* with a 3' mRFP tag and pPJ108 contains cDNA encoding catalytically inactive CLP-1(C371A) (84). Two PCR fragments were used to clone the vector region from PCFJ151, pCFJ151 5' and pCFJ151 3' (Figure 12). The *unc-54::clp-1(C371A)::mRFP* insert was cloned from two fragments; *unc-54::clp-*

1(C371A), which was amplified using primer pair PK1244 and PK1247 and template plasmid pPJ108, and *clp-1::mRFP::3'UTR*, which was amplified using primer pair PK1243 and PK1248 and template plasmid pMC7 (Figure 12). Gibson assembly products were validated by digestion with XhoI. Digestion of correctly assembled pLN1 with XhoI results in the detection of two bands of 10235 bp and 2153 bp. A clone matching this digestion pattern was chosen (Figure 13) and sequenced to confirm its identity.

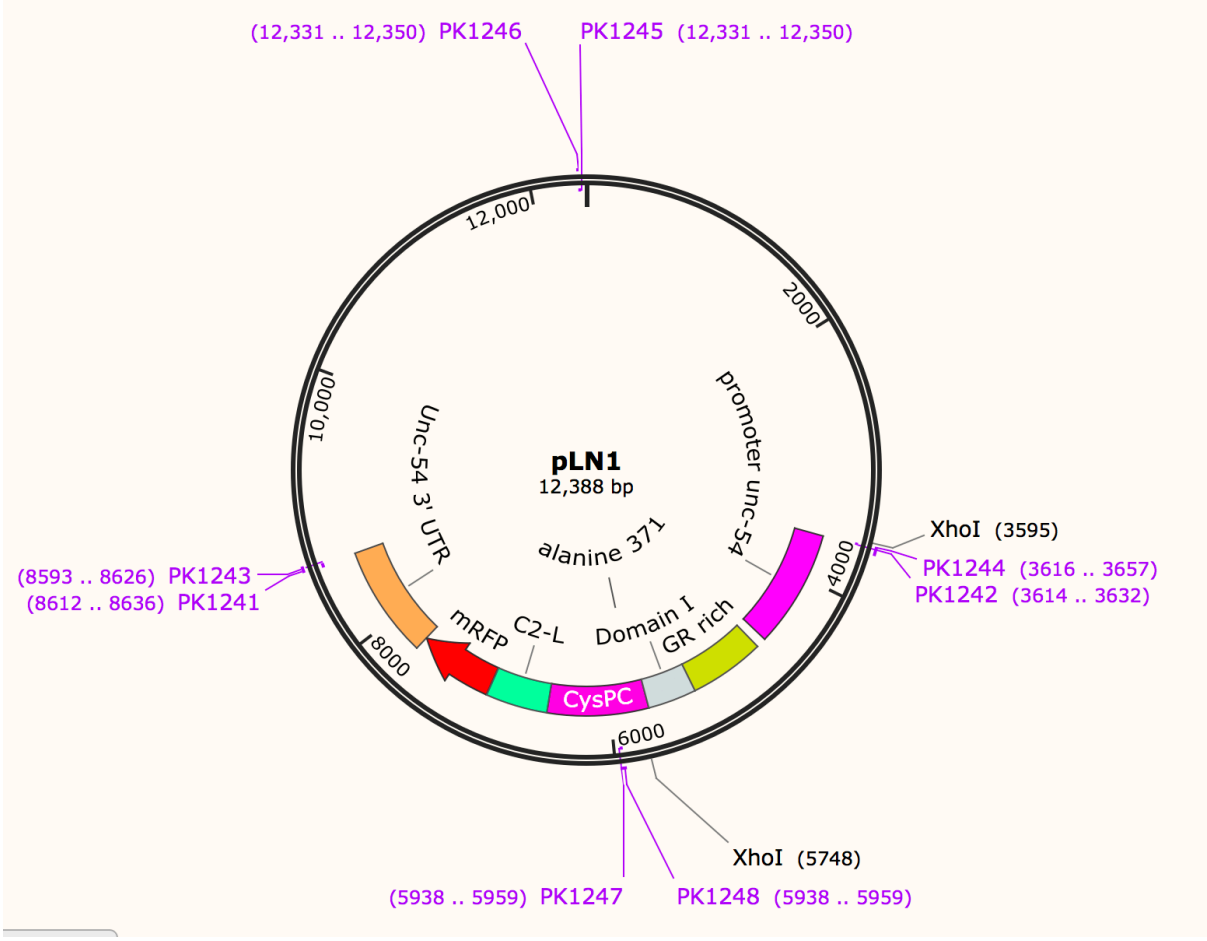


Figure 11. Schematic of the clone pLN1.

pLN1 expresses catalytically inactive CLP-1(C371A)::mRFP under the muscle constitutive *unc-54* promoter (104).

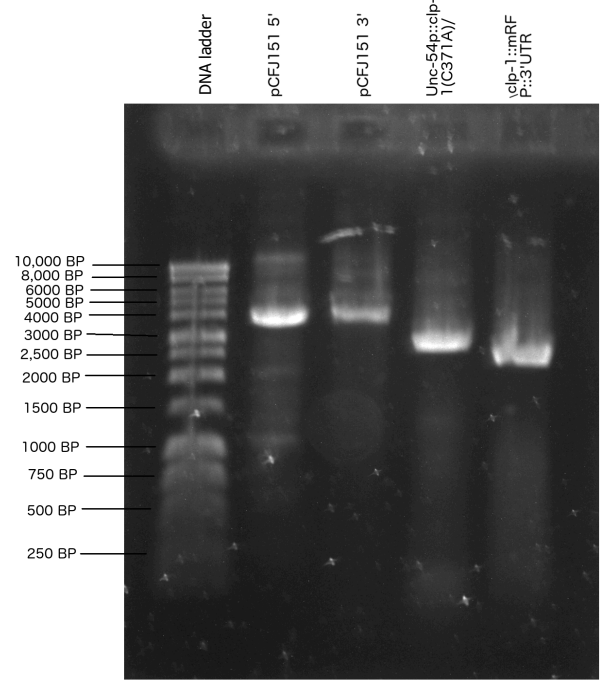


Figure 12. PCR products used to construct pLN1 by Gibson assembly.

1% TBE gel of PCR products used for Gibson assembly of pLN1. pCFJ151 5' is a 3690 bp PCR product generated using primer pair PK1241 and PK1245 on template plasmid pCFJ151, pCFJ151 3' is a 3738 bp PCR product amplified using primer pair PK1242 and PK1246 on pCFJ151 template, *unc-54*::*clp-1*(C371A) is a 2686 bp PCR amplified product using primer pair PK1244 and PK1247 on template plasmid pPJ108 and *clp-1*::mRFP::3'UTR refers to a 2344 bp PCR product using primer pair PK1243 and PK1248 on template pMC7.

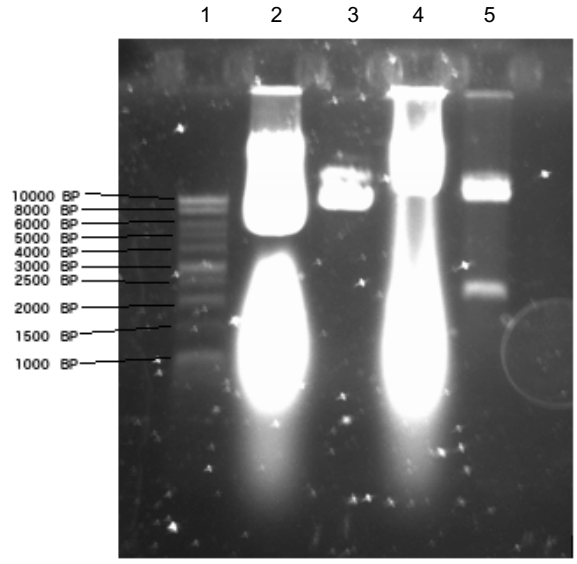


Figure 13. Validation of plasmid pLN1 by XhoI digestion.

Lane 1: DNA ladder. Lane 2 & 4: alkaline lysis minipreps of colonies grown from Gibson assembly product. Lane 3: bands do not correspond with Xho1 digest of pLN1, probable contamination from template plasmids used in PCR reactions for Gibson assembly fragments. Lane 5: bands of 10235 bp and 2153 bp correspond to Xho1 digest of pLN1.

A strain stably expressing (50% transmission) the CLP-1(C371A)::mRFP extrachromosomal array, PK3143, was established by co-microinjecting pLN1 and the marker plasmid pRF4 into 20 adult N2 *C. elegans*. The F1 generation progeny were screened for the roller phenotype and F2 worms were screened for both the roller phenotype and mRFP fluorescence, as discussed in section 3.2. One line was found to express the CLP-1(C371A)::mRFP extrachromosomal array. The extrachromosomal array was integrated into the genome by gamma irradiation. To perform the integration, 50 L4 stage worms expressing CLP-1(C371A)::mRFP were exposed to 48 GY of radiation from a caesium-137 source. 250 F1 worms were picked onto individual seeded plates from irradiated parents and left to lay progeny. Integration of the array was expected to produce a 75% transmission rate amongst the F2 progeny. The five plates containing the highest frequency of worms

expressing CLP-1(C371A)::mRFP as selected by eye were chosen and ten F2 worms from each plate were placed singly on fresh OP50 seeded plates. From these 50 plates, one F3 line showed 100% transmission and the strain PK3113 was established. PK3113 expressed CLP-1(C371A)::mRFP in 100% of animals.

The plasmid clone pLN5 expressing mRFP alone under the control of the constitutive muscle specific promoter of *unc-54* was constructed from pMC7 using site-directed deletion mutagenesis PCR (see section 3.2.1). This clone was microinjected into N2 worms and the strain PK3144 (<30% transmission). This line was integrated using the same methodology used to construct PK3113, yielding strain PK3114.
Exposure to gamma irradiation may result in background mutations in the genome outside of the desired extrachromosomal integration, as well as potential disruption of genes in the locus of where the array was integrated. To mitigate for background mutations, transgenic strains produced by this method were subsequently outcrossed 5x (96).
Genetic mapping was performed on strains PK3113 and PK3114 to determine which chromosome contained the integrated extrachromosomal array. Each strain was crossed with mapping strains PK1003 and PK1009 and the F2 recombinants were assessed (Table 2 and 3). It was shown that in the strain PK3113 the *unc-54::clp-1(C371A)::mRFP* transgene had been integrated into chromosome II, and in PK3114 the *unc-54::mRFP* transgene had been integrated into chromosome X.

F2 genotype	*dpy-5(e61) I*	*bli-2 (e3678) II*	*unc-32(e189) III*	*unc-5(e53) IV*	*dpy-11 (e224) V*	*lon-2(e678) X*
No. of F2 worms	48	33	26	44	19	17
No. of fluorescent worms	25	1	14	22	11	10

Table 2. Mapping the *unc-54::clp-1(C371A)::mRFP* transgene in PK3113.
Fluorescent expression is excluded from F2 worms homozygous for a mutation on chromosome II indicating that the extrachromosomal array carrying the *clp-1(C371A)::mRFP* transgene has been integrated into chromosome II.

F2 genotype	*dpy-5(e61) I*	*bli-2 (e3678) II*	*unc-32(e189) III*	*unc-5(e53) IV*	*dpy-11 (e224) V*	*lon-2(e678) X*
No. of F2 worms	44	38	42	30	32	30
No. of fluorescent worms	33	19	28	16	24	0

Table 3. Mapping the *unc-54::mRFP* transgene in PK3114.
Fluorescent expression is excluded from F2 worms homozygous for a mutation on chromosome X indicating that the extrachromosomal array expressing *mRFP* transgene has been integrated into chromosome X.

3.1.2 Characterisation of integrated *clp-1(C371A)::mRFP* and *mRFP* strains

To assess the suitability of inactive CLP-1(C371A)::mRFP for use in co-immunoprecipitation, the intracellular localisation of CLP-1(C371A)::mRFP was examined (section 3.2) using confocal laser scanning microscopy (CLSM). It was observed that both active CLP-1 and catalytically inactive CLP-1(C371A) were expressed in body wall muscle sarcomeres, localising to m-lines and adjacent to dense bodies (see section 3.2., Figure 32). Therefore, it was anticipated that CLP-1(C371A)::mRFP would have the similar opportunities to bind potential substrates as wildtype CLP-1.

3.1.3 Solubility of CLP-1(C371A)::mRFP in different lysis buffers and immunoprecipitation of CLP-1(C371A)::mRFP

Lysis buffers with different compositions of salts and detergents were tested in their ability to solubilise CLP-1(C371A)::mRFP from whole worm extracts. Ideally the lysis buffer would solubilise the protein of interest so that it could be immunoprecipitated, without being so harsh that it disrupts binding interactions with the protein of interest. When lysates were separated into supernatant and pelleted fractions, CLP-1(C371A)::mRFP was present in the soluble fraction when worms were lysed with either H100 buffer containing 1% Triton X-100 or 0.05% NP-40 and 50 mM KCl.

Lower lysis buffer detergent concentrations are usually recommended to reduce non-specific binding, while preserving protein stability (105). 0.05% NP-40 has been recommended for immunoprecipitation of cytoplasmic proteins from *C. elegans* (106). Therefore, H100 buffer containing 0.05% NP-40 and 50 mM KCl was selected as the solubilisation buffer for co-immunoprecipitation.

Once it was established that CLP-1(C371A)::mRFP was solubilised in 0.05% NP-40 H100 buffer, it was next determined if it could be immunoprecipitated by an RFP nanotrap (ChromoTek). Figure 14 shows the different fractions (input, unbound, bound, flow) resulting from the immunoprecipitation of CLP-1(C371A)::mRFP and mRFP alone using RFP nanotrap (ChromoTek). Both CLP-1(C371A)::mRFP (~110 kDa) and mRFP (27 kDa) were present in the input as expected (Figure 14). The band intensities of CLP-1(C371A)::mRFP and mRFP was slightly diminished but still detectable in unbound fractions, which were sampled after the initial lysate had been incubated with the RFP-nanotrap for 2 hours. This indicated that there was an excess of CLP-1(C371A)::mRFP and mRFP available for binding and retention. The

band intensities of the bound fractions were increased compared to the input and unbound fractions suggesting enrichment by the RFP nanotrap of these proteins (Figure 14). The flow fraction, taken from the final wash of the RFP nanotrap after the 2 hour incubation, showed no presence of the mRFP containing proteins.

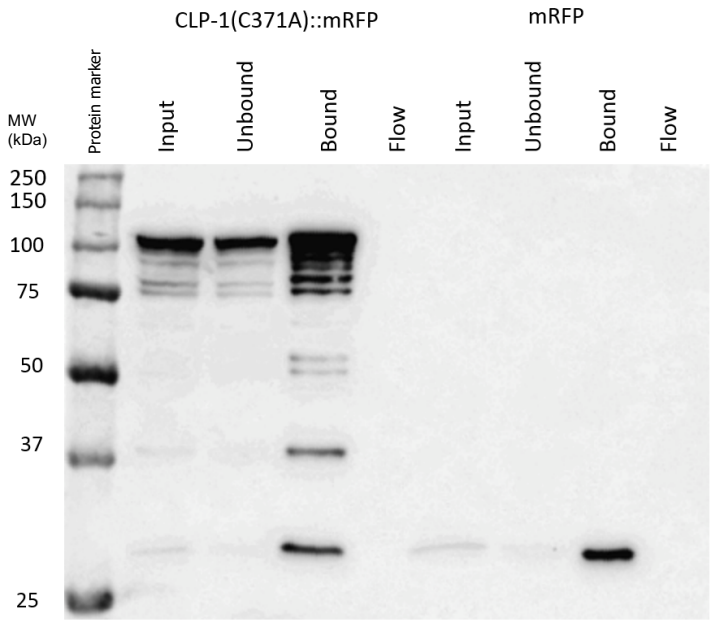


Figure 14. Detection of CLP-1(C371A)::mRFP and mRFP after Immunoprecipitation by RFP-nanotrap.

Western blot of the input, unbound and bound fractions from the immunoprecipitation of CLP-1(C371A)::mRFP (~110 kDa) and mRFP (27 kDa) by RFP-nanotrap from whole worm lysates solubilised in H100 buffer with 0.05% NP-40. The degradation pattern observed of CLP-1(C371A)::mRFP is likely caused by active endogenous CLP-1.

3.1.4 Detecting differences in the proteins co-immunoprecipitated with CLP-1(C371A)::mRFP and mRFP alone

To compare the proteins being co-immunoprecipitated by CLP-1(C371A)::mRFP and mRFP alone, the input and bound fractions were electrophoresed on a 10% SDS-PAGE gel and silver-stained (Figure 15). Several bands appeared to be enriched in the bound fraction of CLP-1(C371A)::mRFP when compared to the bound fraction of mRFP alone (Figure 15).

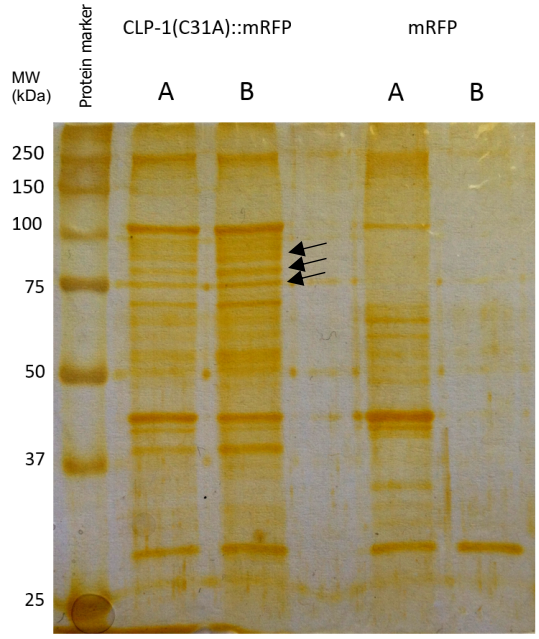


Figure 15. Silver stain of the immunoprecipitation of CLP-1(C371A)::mRFP and mRFP from whole worm lysates.

Some protein bands are enriched in the bound fraction of CLP-1(C371A)::mRFP (B) compared to the input lane (A). There are also protein bands that are enriched in the bound fraction of CLP-1(C371A)::mRFP compared to the bound fraction of mRFP (B). Arrowheads: enriched protein bands.

Co-immunoprecipitation was replicated three times using three independent worm preps grown at 20°C. Lysis of these worm preps and co-immunoprecipitation of CLP-1(C371A)::mRFP and mRFP alone were performed and the bound fraction sent to the faculty of biomedical sciences proteomics facility for mass-spectrometry analysis.

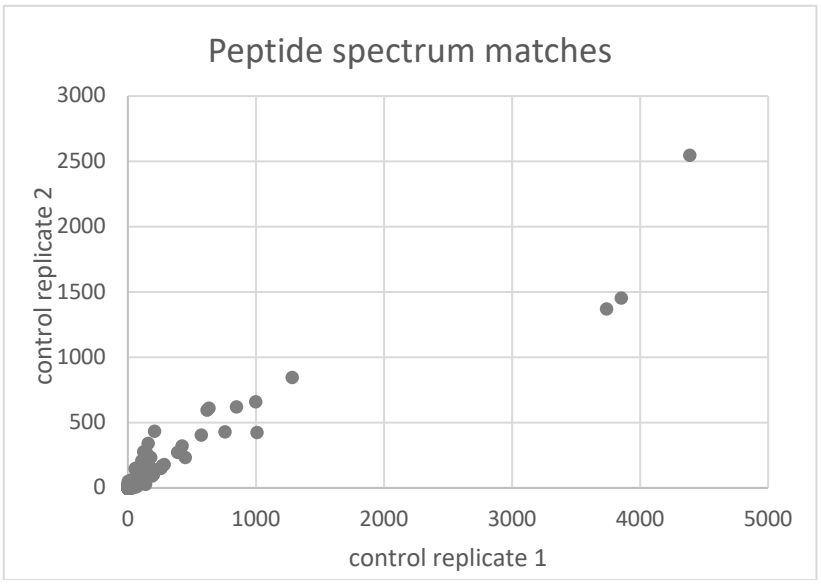
3.1.5 Proteomic analysis

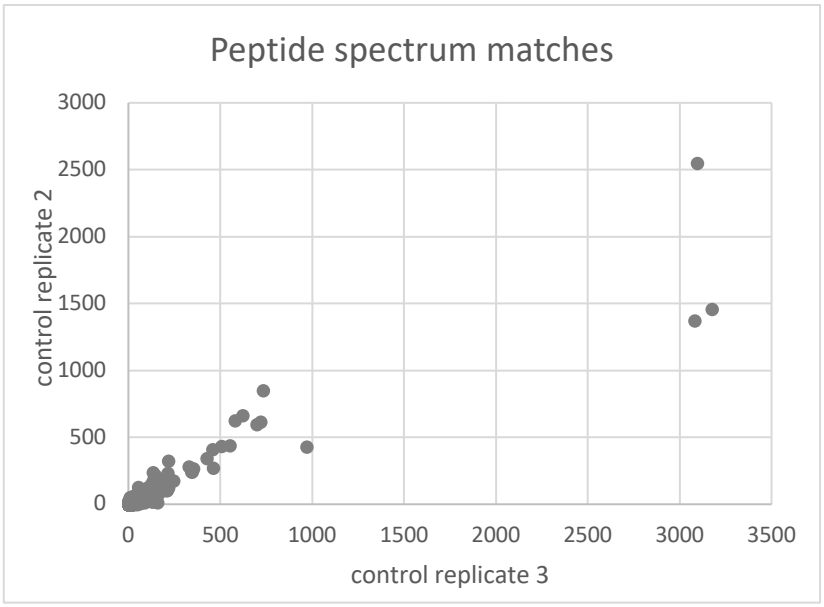
Samples containing proteins co-immunoprecipitating with CLP-1(C371A)::mRFP and mRFP alone were electrophoresed on polyacrylamide gels. Gel lanes were cut into slices and digested with trypsin before performing high performance liquid chromatography and Orbitrap Velos (Thermo Scientific) tandem mass spectrometry analysis (107). Mass spectra collected from the orbitrap were subjected to a SEQUEST search against the theoretical mass spectra of all *C. elegans* proteins. A peptide spectrum match (PSM) occurs when the actual mass spectra is similar to the theoretical mass spectra of a *C. elegans* protein. To increase confidence in the identification of a PSM, the search was performed at False discovery rate (FDR) of 1%. False discovery rate is assessed by the number of peptide spectrum matches observed in a control set of theoretical proteins, in this case the complete *C. elegans* proteome with the sequences reversed.

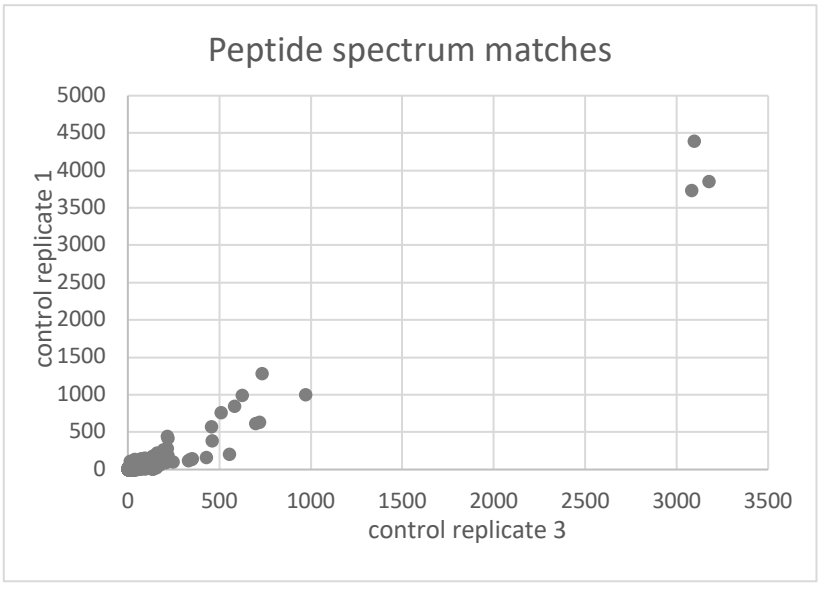
3.1.5.1 Biological replicates show significant similarity

Before performing statistical analysis on the proteomic replicate data, reproducibility between the replicates was analysed to ensure that subsequent analysis was meaningful. Pairwise comparisons between all replicates were performed; Spearman rank correlation and Pearson correlation was calculated for each pair of proteomic data (Table 4). Spearman correlation measures monotonicity of a relationship between two variables, and Pearson correlation measures linear correlation between two variables. A minimum of 0.77 Spearman rank correlation and 0.94 Pearson correlation was observed between biological replicates, indicating correlation between biological replicates comparable to values found in other proteomic studies

(108). As can be seen in scatter plots of the PSM count of proteins in each replicate pair (Figure 16) there are several proteins with large PSM counts (>1000), the presence of these proteins with large PSM count may affect Pearson's correlation coefficient which is sensitive to outliers (109). Spearman's rank correlation is not affected by the presence of these large PSM count outliers as it depends on the rank of the proteins not their PSM count, and so it is a more valid measure of correlation between these replicates. The pairwise comparisons showed that there was sufficient reproducibility between replicates to use them for statistical analysis when determining which proteins were enriched in the CLP-1::mRFP co-IP vs the control mRFP alone co-IP.


Peptide spectrum matches
control replicate 2
control replicate 1
0
500
1000
1500
2000
2500
3000
0
1000
2000
3000
4000
5000


Peptide spectrum matches
control replicate 2
control replicate 3
0
500
1000
1500
2000
2500
3000
0
500
1000
1500
2000
2500
3000
3500


Peptide spectrum matches
control replicate 1
control replicate 3
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
0
500
1000
1500
2000
2500
3000
3500

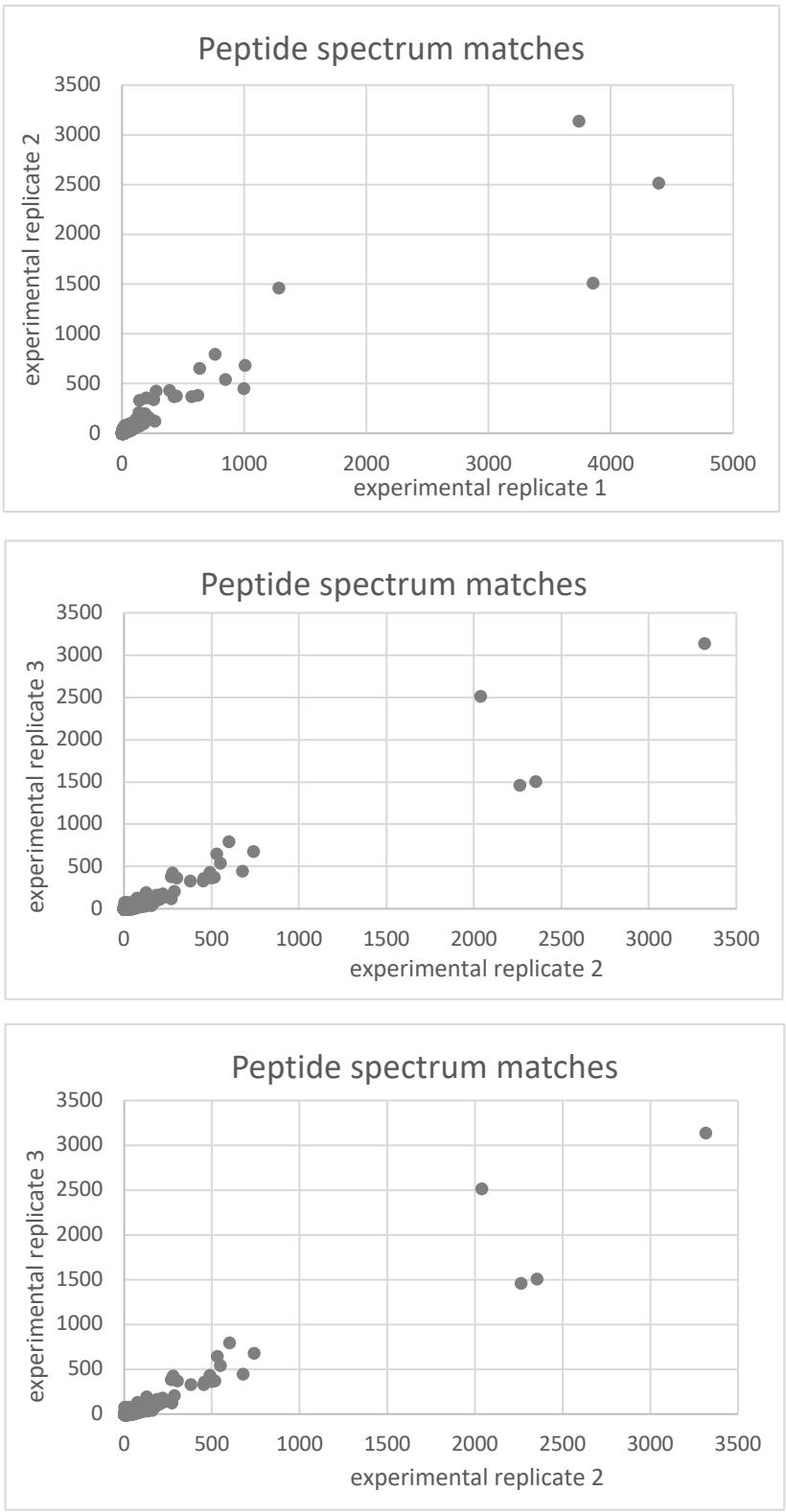


Figure 16: Scatter plots of peptide spectrum match of replicate pairs. These plots show pairwise correlation between control replicates and between experimental replicates. Each replicate contains several proteins with large PSM counts (>1000 PSM).

Spearman	1	2	3
1	X	0.85*	0.78*
2	0.83*	X	0.82*
3	0.77*	0.79*	X
Pearson	1	2	3
1	X	0.95*	0.97*
2	0.95*	X	0.97*
3	0.98*	0.94*	X
Control			
Experimental			

Table 4. Pairwise Spearman and Pearson correlation of PSM data from the triplicate biological replicates of CLP-1::(C371A)::mRFP and mRFP alone. For both control and experimental co-IPs, the minimum Spearman rank correlation observed between biological replicates was 0.77. These values were >0.7 which is acceptable for this type of experiment (108). *all p-values < 0.01 for Pearson and Spearman correlation for each pairwise comparison. n for each pairwise comparison: control replicate pair 1 & 2; n = 1277, control replicate pair 2 & 3; n = 1267, control replicate pair 1 & 3; n = 1540, experimental replicate pair 1 & 2; n = 1607, experimental replicate pair 2 & 3; n = 1676, experimental replicate pair 1 & 3; n = 1709.

3.1.5.2 Proteins enriched with CLP-1

To identify proteins that form complexes with CLP-1(C371A)::mRFP a statistical analysis comparing the PSM counts of proteins identified in the mRFP co-IP vs the CLP-1(C371A)::mRFP co-IP was performed. Biological significance was measured by the fold change of the PSM count, and statistical significance was measured by 2-sample Students t-test of PSM counts of the CLP-1(C371A)::mRFP co-IP and mRFP co-IP triplicates. Cut off values were used to narrow the number of positive interactors and reduce identification of false positives (110).

SAINTexpress (significance analysis of interactome) is software used for probabilistic scoring of protein-protein interaction data using a iterated conditional

mode Markov random field model (111). SAINTexpress was used to calculate the fold change (FC-a); when the PSM count was zero, it was replaced by 0.1 to avoid division by zero and enable statistical analysis (111). To assess the statistical significance of PSM counts identified in the experimental group of replicates versus the control group, a Students t-test with no assumption of equal variance was performed.

Cut-off values of 2.5x fold change and p-value <0.05 were chosen to reduce likelihood of false positives. 57 proteins in total were identified that met these criteria (Table 5). This set of proteins was used to perform enrichment analysis and were further examined as potential CLP-1 substrates. As might be predicted, CLP-1 showed the greatest fold-change of an average ~63x between the experimental and control co-IPs (Table 5).

A volcano plot, plotted using Microsoft Excel, showing the distribution of significant hits from this experiment amongst all proteins identified in both runs was plotted (Figure 17). The grouping of proteins around CLP-1 and absence of this grouping in the control side (negative x-axis) indicates the presence of CLP-1 enriched partners and not mRFP specific enriched partners.

3.1.5.3 Gene enrichment analysis and bioinformatics tool WormGRAB

To assist with this analysis, I developed a bioinformatics tool named WormGRAB . When WormGRAB is provided a list of *C. elegans* prey proteins and a single bait protein input, it requests and processes information from the representational state transfer application programming interface (REST API) of Wormbase (79). The

information generated includes; available reagents including worm strains, antibodies and plasmids, as well as GO terms, expression and phenotypic data (79). WormGRAB (Figure 18, 19) also uses PubMed's API to search for instances of co-referencing between the bait and each prey protein, and then also searches for co-referencing between their homologues. If a co-reference is identified, the unique PubMed IDs (PMIDs) for the relevant papers are retrieved (112). WormGRAB is implemented using a node.js http server that makes asynchronous calls to the APIs Wormbase and PubMed, implemented using a callback pattern (Figure 20). WormGRAB is free to access on all modern browsers at http://wormgrab.herokuapp.com/. The code repository for WormGRAB is available online at https://github.com/lauri3new/wormgrab. For phenotype enrichment analysis the online Wormbase tool was used (Table 6) (113).

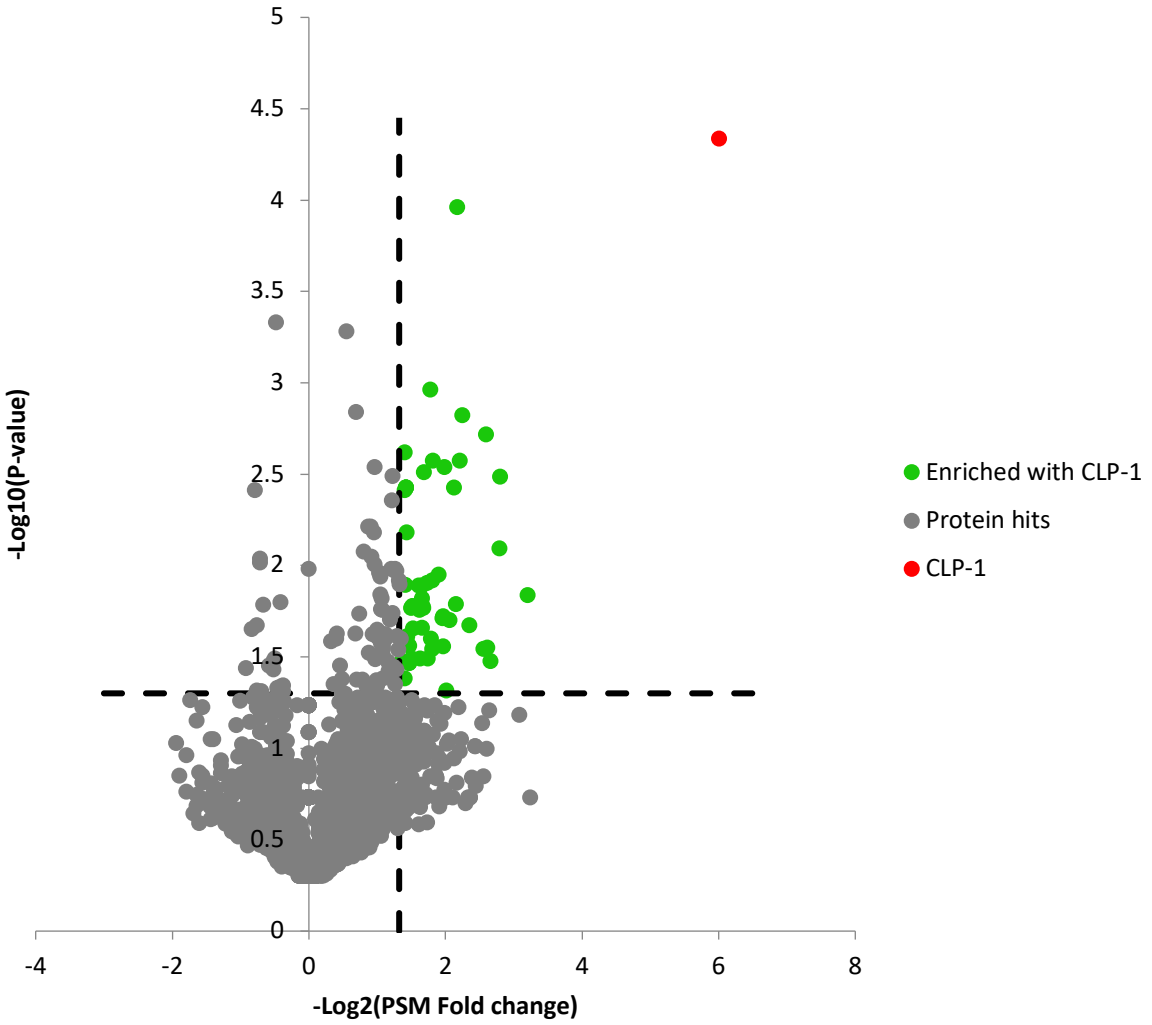


Figure 17. Volcano plot of proteins enriched in CLP-1::mRFP IP vs mRFP alone IP.

Proteins (green) clustered with CLP-1 (red) are those that are enriched in the CLP-1::mRFP co-IP compared with mRFP alone co-IP.

Accession	Description*	Fold change
P34308	clp-1 encodes a calpain homolog.	63.04
Q19132	F07A11.4 is an ortholog of human USP19 (ubiquitin specific peptidase 19).	8.01
Q9N5G1	dhs-15 encodes a member of the short chain dehydrogenase/reductase family.	6.89
O01884	coq-6 encodes a putative flavin-dependent aromatic-ring monooxygenase.	6.36
P90904	K02B12.7 is an ortholog of human ArfGAP1 (ADP ribosylation factor GTPase activating protein 1)	5.55
Q3HKC4	pcyt-1 encodes a lipid-activated CTP:phosphocholine cytidylyltransferase (CCT).	5.3
Q23179	Microarray and RNA sequencing studies indicate that W05H9.1 is affected by Tryptophan…	5.86
Q965R9	Y97E10AR.6 is an ortholog of human BSDC1 (BSD domain containing 1).	5.21
O76566	unc-96 encodes a novel protein with no recognizable domains.	4.43
H2L0N0	Microarray, proteomic, and RNA sequencing studies indicate that unc-132 is affected by cyc-1…	4.63
P91270	F25B4.7 is an ortholog of human SLC25A4 (solute carrier family 25 member 4).	4.63
Q9TZH8	T10B11.6 is an ortholog of human TMEM53 (transmembrane protein 53).	4.44
O17766	lipl-7 is an ortholog of members of the human Lipases family including LIPF.	4.05
Q9XUR4	T05F1.2 encodes a novel protein.	4.27
Q7JPQ6	RNA sequencing and microarray studies indicate that B0361.3 is affected by jmjd-3.1…	3.86
O61196	F37C4.6 is an ortholog of human PYROXD2 (pyridine nucleotide-disulphide oxidoreductase domain 2).	3.48
G5EDP6	fmo-1 encodes a flavin-containing monoxygenase homologous to human FMO1, FMO2, and FMO3.	3.82
P34666	coq-5 encodes a putative 2-hexaprenyl-6-methoxy-1,4-benzoquinone methyltransferase.	3.56
Q93934	R07H5.8 is an ortholog of human ADK (adenosine kinase).	3.67
Q7K6X4	npp-16 is an ortholog of human Nup50 (nucleoporin 50).	3.53
Q18486	coq-8 encodes a putative protein kinase orthologous to ABC1 (COQ8) from S. cerevisiae and UbiB from E. coli.	3.63
Q9U3Q6	ugt-22 is an ortholog of human UGT3A2 (UDP glycosyltransferase family 3 member A2) and UGT3A1.	3.53
H2KYN0	vps-45 is an ortholog of human Vps45 (vacuolar protein sorting 45).	3.2
Q9NET7	Microarray and RNA sequencing studies indicate that Y39B6A.10 is affected by Hydrogen sulfide.	3.35
Q95QH4	F32A5.8 is an ortholog of human Wwox (WW domain containing oxidoreductase).	3.16
Q9GYR8	Microarray and RNA sequencing studies indicate that C23G10.5 is affected by lin-15B, pgl-1, pgl-3, glh-4.	3.36
G5ED43	fmo-5 encodes a flavin-containing monoxygenase homologous to human FMO1, FMO2, and FMO3.	3
G5EBU3	zmp-3 is an ortholog of human Mmp1 (matrix metallopeptidase 1) and members of the MMP (M10 matrix metallopeptidases).	3.19
Q9U1Q2	gpi-1 encodes two isoforms of a putative glucose 6-phosphate isomerase orthologous to human GPI.	3.16
I2HA81	Microarray studies indicate that F15B9.10 is affected by Humic substances…	3.02
Q9XVI1	RNA sequencing and microarray studies indicate that T23G11.1 is affected by Ethanol…	3.02
Q9N4G7	Y71F9B.2 is an ortholog of human TAMM41 (TAM41 mitochondrial translocator assembly and maintenance).	2.98
Q09EE6	RNA sequencing and microarray studies indicate that K07A1.17 is affected by Rotenone…	2.96
G5EF84	RNA sequencing and microarray studies indicate that gip-2 is affected by tatn-1…	2.95
H2L0K4	Microarray, RNA sequencing, and tiling array studies indicate that Y45G5AL.1 is affected by hlh-30…	2.85
Q09289	rpn-13 is an ortholog of human ADRM1 (adhesion regulating molecule 1).	2.97
H1ZUX7	-	2.92
Q9UAT3	mct-2 is an ortholog of members of the human SLC (Solute carriers) family including SLC16A3.	2.76
Q9N3C2	Microarray and RNA sequencing studies indicate that Y54F10AR.2 is affected by cyc-1…	2.78
Q9U2X0	prmt-1 encodes a type I protein arginine methyltransferase.	2.64
Q9U2Y2	cdkr-3 is an ortholog of human CDK5RAP3 (CDK5 regulatory subunit associated protein 3).	2.72
P54813	ymel-1 encodes an ortholog of human YME1L1.	2.52
O17397	F52H2.6 is orthologous to the human gene 3BETA-HYDROXYSTEROL DELTA 24 REDUCTASE.	2.52
Q93595	F26A3.1 is an ortholog of human LDAH (lipid droplet associated hydrolase).	2.57
H2KMK6	Y38E10A.22 is an ortholog of human HSPBP1 (HSPA (Hsp70) binding protein 1).	2.64
W6RY92	T12B3.4 encodes one of three *C. elegans* proteins orthologous to members of the MOB (Mps One Binder).	2.64
Q8MXQ7	Y92H12BL.1 is an ortholog of human CDKAL1 (CDK5 regulatory subunit associated protein 1 like 1).	2.52
O16487	Microarray, RNA sequencing, tiling array, and proteomic studies indicate that B0238.11 is affected by cgh-1…	2.64
A4UVL1	Microarray and RNA sequencing studies indicate that D1086.17 is affected by nhr-114…	2.64
P30639	lnkn-1 encodes a conserved, single-pass type I transmembrane glycoprotein.	2.64
O02345	ZK896.9 encodes a Golgi apparatus transporter for UDP-glucose, UDP-galactose, UDP-N-acetylglucosamine	2.64
P46502	rpt-3 encodes a triple A ATPase subunit of the 26S proteasome's 19S regulatory particle (RP) base subcomplex.	2.67
Q2HQK3	Y62E10A.20 is an ortholog of human C2orf47.	2.52
Q7Z2A5	mop-25.1 encodes a protein orthologous to fission yeast Mo25p.	2.62
O16266	Microarray and RNA sequencing studies indicate that F40A3.6 is affected by Methylmercuric chloride…	2.5
Q21557	dlc-2 encodes a putative dynein light chain 1.	2.63
Q21443	lmn-1 encodes the sole *C. elegans* nuclear lamin.	2.54
Q17439	B0035.15 is an ortholog of human NDUFAF4 (NADH:ubiquinone oxidoreductase complex assembly factor 4).	2.51

Table 5. 58 proteins that are most enriched in the CLP-1(C371A)::mRFP co-immunoprecipitation.

As expected the bait protein CLP-1 is the most enriched protein in the experimental co-IP. Fold change represents the average PSM ratio in experimental to control co-IP experiments across all replicates and was calculated using SAINTexpress software.*Description directly abridged from Wormbase description for the gene (71).

3.1.5.4 Gene enrichment analysis of proteins co-immunoprecipitating with CLP-1

Phenotype					
Term	Expected	Observed	Enrichment Fold Change	P-value	Q-value
multiple nuclei oocyte	0.23	3	13	9.30E-05	0.02
muscle system morphology variant	0.97	5	5.2	0.00046	0.051

Table 6. Phenotype enrichment analysis.
Phenotype enrichment analysis shows proteins important for proper muscle morphology that are enriched by co-immunoprecipitation with CLP-1.

When CLP-1 is inappropriately activated, it could lead to dysregulated proteolysis of target substrates. In the case of substrates in body wall muscle, the resulting loss of muscle integrity could lead to paralysis. Under standard conditions CLP-1 might maintain proper muscle morphology by regulating the turnover of these proteins (10,77).

It is possible that CLP-1 might also have a role in oocyte development given the enrichment of proteins with mutants producing multiple nuclei oocyte phenotypes, previously in starfish calpain was shown to be involved in meiosis reinitiation in oocytes (114).

3.1.5.5 Potential CLP-1 substrates

To assess and narrow which proteins identified from the co-IP MS experiment would be selected for further validation, 3 subsets of the 57 proteins that met the cut off criteria were examined: the proteins with the greatest fold change (Table 7), the proteins that give rise to muscle disorganisation phenotypes when disrupted through RNAi or knockout (Table 8), and the proteins which have orthologues that have been previously identified as possible calpain substrates are displayed in Table 9.

Accession	Gene	Description*	Fold change
Q19132	*F07A11.4*	F07A11.4 is an ortholog of human USP19 (ubiquitin specific peptidase 19).	9.21
Q9N5G1	*dhs-15*	dhs-15 encodes a member of the short chain dehydrogenase/reductase family.	6.97
O01884	*coq-6*	coq-6 encodes a putative flavin-dependent aromatic-ring monooxygenase, homologous to yeast COQ6.	6.92
P90904	*K02B12.7*	K02B12.7 is an ortholog of human ArfGAP1 (ADP ribosylation factor GTPase activating protein 1).	6.35

Table 7. Proteins detected by mass spectrometry (1% FDR) ranked by fold change.
Fold change represents PSM ratio in experimental to control IP and was calculated using SAINTexpress. *Description abridged from Wormbase description (71).

The proteins that were most enriched in the co-IP were FO7A11.4, DHS-15, COQ-6 and K02B12 (Table 7). FO7A11.4 encodes an ortholog of human ubiquitin specific peptidase 19 (USP19). Interestingly, inactivation of USP19 was shown to decrease muscle atrophy in mice (115). COQ-6 monooxygenase is required for coenzyme q10 synthesis. The electron chain component coenzyme q10 is primarily found in mitochondria and its deficiency has been associated with several diseases including Parkinson's disease (116). It has also been investigated as a potential therapy for Duchenne muscular dystrophy due to it being an antioxidant (117).

Accession	Name	Ortholog	Description*	Fold change
Q3HKC4	*pcyt-1*	Phosphocholine Cytidylyltransferase	pcyt-1 encodes a lipid-activated CTP:phosphocholine cytidylyltransferase (CCT).	6.13
Q9U2Y2	*cdkr-3*	CDK5 Regulation associated protein	cdkr-3 is an ortholog of human CDK5RAP3 (CDK5 regulatory subunit associated protein 3); cdkr-3 is expressed in the pharynx, spermatheca, nervous system, and the intestine.	2.83
Q21443	*lmn-1*	nuclear Lamin	lmn-1 encodes the sole *C. elegans* nuclear lamin.	2.62

Table 8. Targets identified in proteomic analysis that have orthologues that are possible calpain substrates.
*Description abridged from Wormbase description (71).

pcyt-1 encodes a lipid-activated CTP:phosphocholine cytidylyltransferase, which is a rate-limiting enzyme in biosynthesis of phosphatidylcholine. It was previously shown that oxidised low density lipids inhibited CTP:phosphocholine cytidylyltransferase via calpain degradation (118). Levels of phosphatidylcholine in human skeletal muscle are linked to insulin sensitivity and exercise response (119).

CDKR-3 is an ortholog of human CDK5 regulatory subunit associated protein 3, although there is no evidence that calpain cleaves CDK5RAP3, calpain has been shown to digest the CDK5 regulatory protein p35 to produce the protein p25 and this cleavage is associated with neuronal apoptosis (120).

lmn-1 encodes a nuclear lamin, nuclear lamins are possible calpain substrates in starfish (114). Nuclear lamins are fibrous proteins that are important for transcriptional regulation and structural stability of the nucleus. Degradation of nuclear lamin is a feature of apoptotic cell death (121). Mutations in lamins can result

in laminopathies with symptoms of muscular dystrophy, neuropathy and premature aging (122).

Accession	Gene	Description*	Muscle expression	Locomotion reduced/ Variant**	Apoptosis variant**	Body wall muscle defective/ variant**	Fold change
Q9TZH8	*T10B11.6*	T10B11.6 is an ortholog of human TMEM53 (transmembrane protein 53).	Yes	-	-	Yes	4.51
Q3HKC4	*pcyt-1*	pcyt-1 encodes a lipid-activated CTP:phosphocholine cytidylyltransferase (CCT	-	Yes	Yes	Yes	6.13
O76566	*unc-96*	unc-96 was identified in screens for mutations that disrupt body wall muscle cell structure.	Yes	Yes	-	Yes	5.10

Table 9. Targets identified that produce muscle variant phenotypes in RNAi screens.
Although not much is known about transmembrane protein 53's function, it was found that a knockdown of NET4/TMEM53 in cell type caused premature senescence in primary fibroblasts (123). *unc-96* mutants have decreased locomotion and myofibrillar organisation. UNC-96 is required for correct assembly and maintenance of *C. elegans* body wall muscle sarcomeres (124). *Description abridged from Wormbase description (71). **Phenotype information from RNAi screening (79).

LMN-1 and TMEM-53 are nuclear and mitochondrial proteins, respectively. CLP-1::mRFP was not detected in either of these structures (section 3.2) by LSCM. It might be the case that these are artefacts produced by the extraction process, or that they bound together in disrupted or apoptotic cells. Although the screen yielded a number of possible substrates, it remains necessary to validate these targets as calpain substrates. Some of the targets identified will likely be false positives (although steps were taken in the analysis such as using a 1% FDR) and not all interacting partners might be substrates, the sources of false positives and possible

validation experiments are further discussed in section 4.1. However, due to time limitation these validation experiments were not able to be completed.

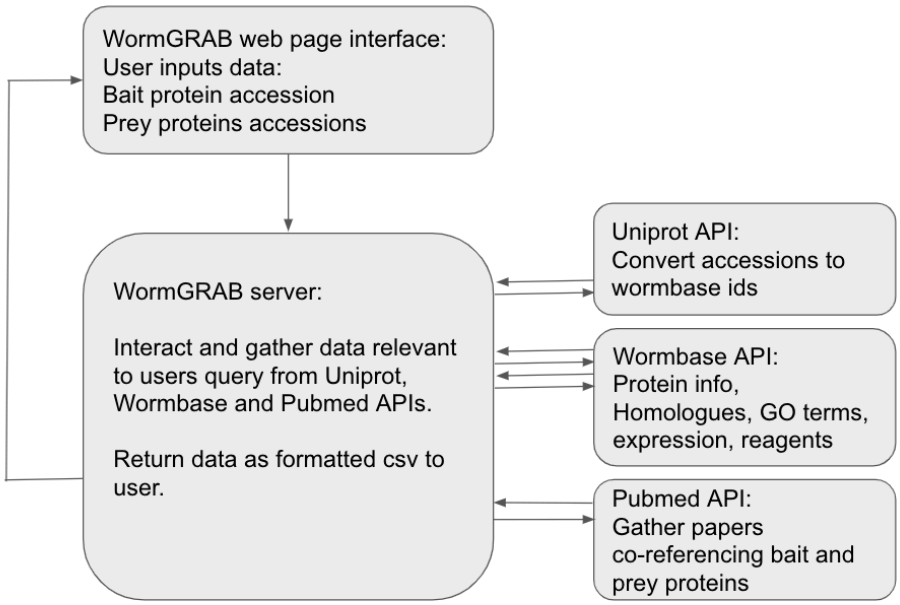


Figure 18: WormGRAB protein and PPI data collection tool.

WormGRAB gathers protein and gene data and literature references from Wormbase and Pubmed APIs about bait and prey proteins inputted by users via a web interface. Arrows correspond to http requests.

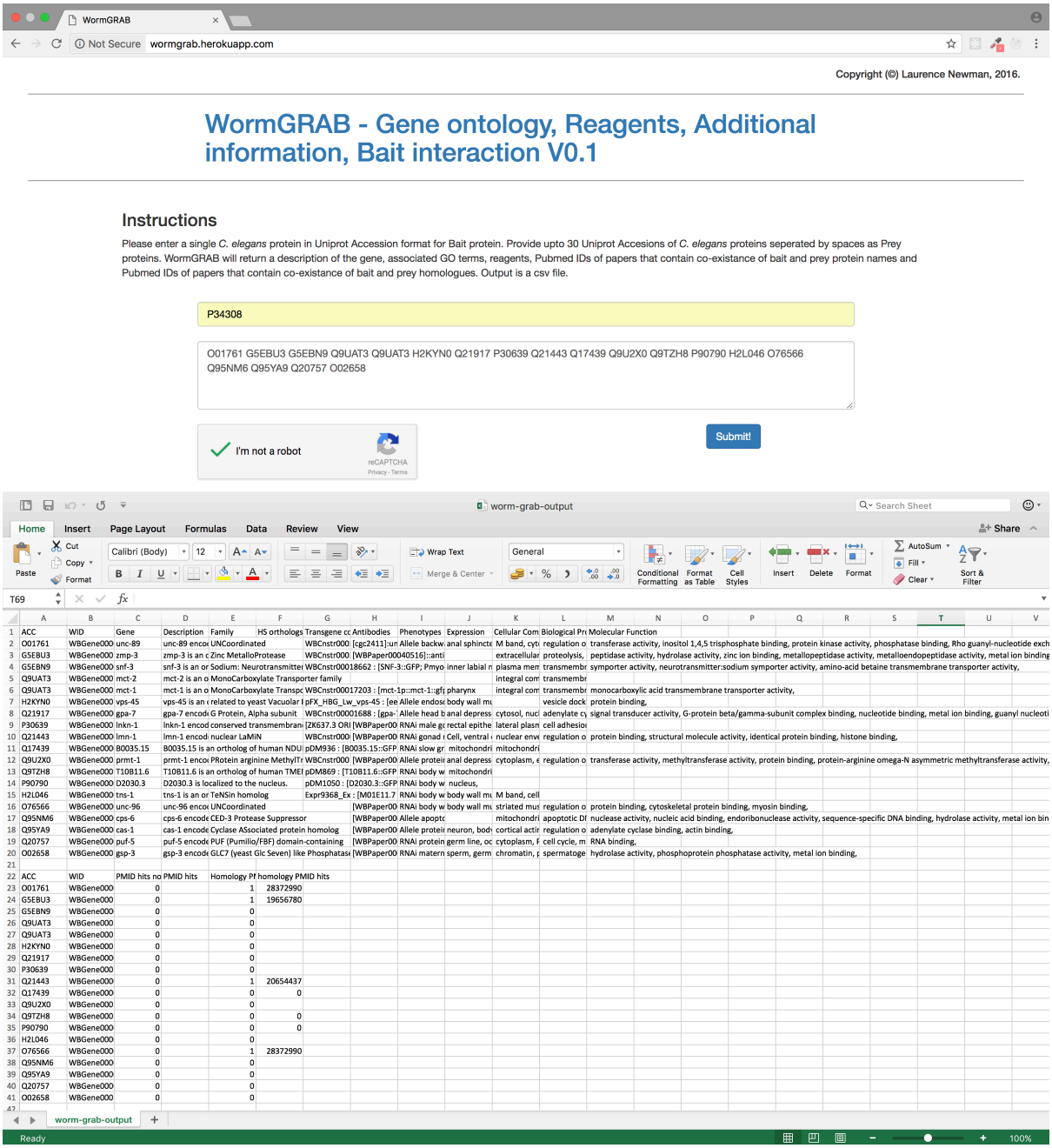


Figure 19: WormGRAB interface and output.

Using Uniprot accession identifiers as input, WormGRAB interacts with Uniprot, Wormbase and PubMed API's to gather information that can be used to analyse co-IP MS experiments.

```
// get concise_description from wormbase about wormbase ID, y pushes to outarr[y] in 2d array, callback getReagents
function getDesc(WID, y) {
    var reqURL = 'http://api.wormbase.org/rest/widget/gene/' + WID + '/overview';
    request.get({
        url: reqURL,
        json: true,
        headers: {
            'Content-Type': 'Content-Type:application/json'
        }
    }, function(error, response, body) {
        if (error) {
            console.log(error);
        } else {
            outarr[y].push(body.fields.name.data.label);
            outarr[y].push(body.fields.concise_description.data.text);
            var family = "";
            if (body.fields.gene_class.data !== null) {
            searcharr[y+1].push("(" + body.fields.gene_class.data.description.split(" ").join("+AND+").replace(/\n/g,"") + ")");
            family += body.fields.gene_class.data.description;
            }
            outarr[y].push(family);
            getHomology(WID, y);
        }
    })
}
```

Figure 20. Node.js code snippet of WormGRAB.

A code snippet of JavaScript code from WormGRAB, showing a http request to API call and http response processing. Multiple API calls are combined by the callback pattern to build up an array of arrays of data that is outputted as a tab separated text file for the user.

3.1.6 Candidate gene analysis

Vinculin, beta-integrin (PAT-3) and intermediate filaments were investigated as potential binding partners of catalytically inactive CLP-1(C371A). Vinculin, beta-integrin (PAT-3) and intermediate filaments are present in the sarcomeres of *C. elegans* body wall muscle cells and orthologues have been reported to be substrates for typical calpains (13,14,15). The crosslinking protein vinculin is found in the integrin attachment complexes of the sarcomere; M-lines, dense bodies and attachment plaques (see Figure 9). Beta-integrin plays a structurally important role in the sarcomeres of body wall muscle cells, anchoring dense bodies and m-lines to the sarcolemma. Intermediate filaments provide the link between the hypodermis and epidermis (42).

To test for potential interaction between CLP-1 and these target proteins, co-IP by RFP nanotrap fractions of CLP-1(C371A)::mRFP and mRFP alone were analysed by SDS-PAGE and western blotting using mouse monoclonal antibodies MH4, MH24 and MH25. MH4 recognises an intermediate filament subunit (125), MH24 recognises vinculin (75) and MH25 recognises beta-integrin (126). It was expected that a binding partner would be detected against the CLP-1(C371A)::mRFP, but not mRFP alone bound fractions of the respective co-IPs. Candidates were lysed in H100 buffer containing low detergent 0.05% NP-40, high detergent (1% Triton X-100) and high salt (300mM KCL) and the solubility and co-immunoprecipitation with CLP-1(C371A)::mRFP is summarised in Table 10.

Condition	Soluble in 0.05% NP-40, 100 mM KCL	Soluble in 0.05% NP-40, 300 mM KCL	Soluble in 1% Triton X-100, 100 mM KCL	Co-immunoprecipitated w/ CLP-1(C371A)::mRFP
PAT-3	X	X	✓	✓
IFA-1	X	X	X	X
Vinculin	✓	✓	✓	X

Table 10. Lysis buffers candidate proteins were tested in.
PAT-3 was soluble when lysed in h100 buffer containing 1% Triton X-100, and co-immunoprecipitated with CLP-1(C371A)::mRFP under these conditions. Vinculin was soluble in all lysis buffers tested but did not co-immunoprecipitate with CLP-1(C371A)::mRFP. IFA-1 was insoluble in all buffers tested.

3.1.6.1 Vinculin and Intermediate filaments

Vinculin did not co-immunoprecipitate with CLP-1(C371A)::mRFP (Figure 21), indicating that vinculin is not a binding partner of CLP-1 and is unlikely to be a substrate, although it could be that vinculin is a transient or low-affinity binding partner and therefore not be detected in a co-IP experiment. Intermediate filament A (IFA) was not soluble in H100 buffer (Figure 22) or in high detergent or high salt conditions. A harsher lysis buffer could be used to solubilise IFA however it is unlikely to preserve an interaction with CLP-1.

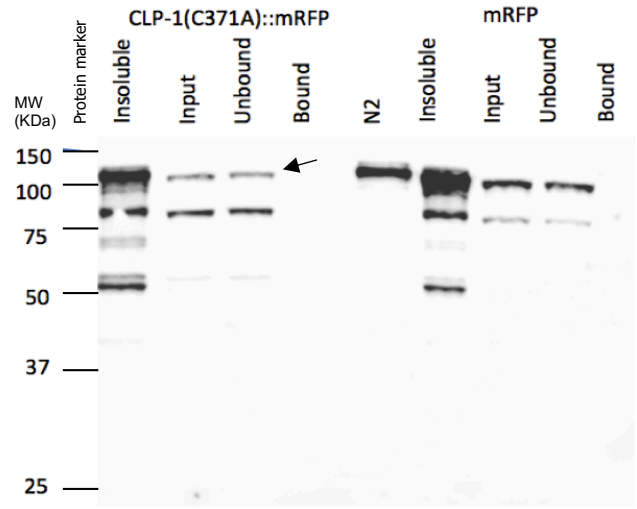


Figure 21. Vinculin (~108Kda) does not co-immunoprecipitate with CLP-1(C371A)::mRFP.

Western blot of co-IP fractions stained for vinculin using MH24 primary antibody. Vinculin is the band at approximately 108 kDa. Vinculin is not present in the bound fraction of CLP-1(C371A)::mRFP. The band labelled N2 is a prep of 80 whole N2 worms. Arrowhead: vinculin.

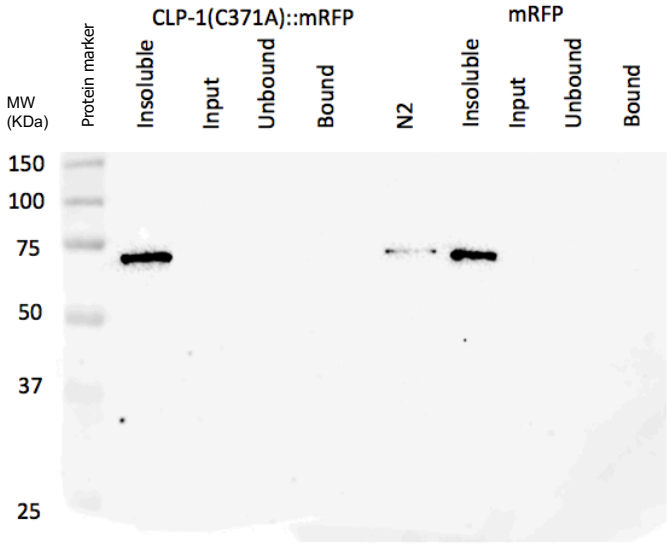


Figure 22. Intermediate filament A is insoluble in H100 lysis buffer.

Western blot of co-IP fractions immunostained for IFA using MH4 primary antibody. IFA (~70Kda) is insoluble in H100 lysis buffer. IFA did not solubilise under increased detergent or salt conditions. The band labelled N2 is a prep of 80 whole N2 worms.

3.1.6.2 PAT-3 beta-integrin co-immunoprecipitates with CLP-1

Beta-integrin was present in the insoluble fraction of the lysate only when frozen worm pellets were lysed in 0.05% NP-40 H100 buffer. When worm pellets were lysed in 1% Triton X-100 H100 lysis buffer, beta-integrin was present in the soluble input fraction. Beta-integrin was present in the bound fraction of the co-IP of CLP-1(C371A):: mRFP but not of the co-IP of mRFP alone. This suggests that beta-integrin forms a protein complex with CLP-1(C371A) (Figure 23). PAT-3 was not detected in the mass-spec analysis. Given that PAT-3 is not solubilised by the reduced detergent used in the lysis buffer for the proteomic experiment, it could explain why it was not detected. Further validation experiments are required to confirm if beta-integrin is a catalytic substrate of CLP-1.

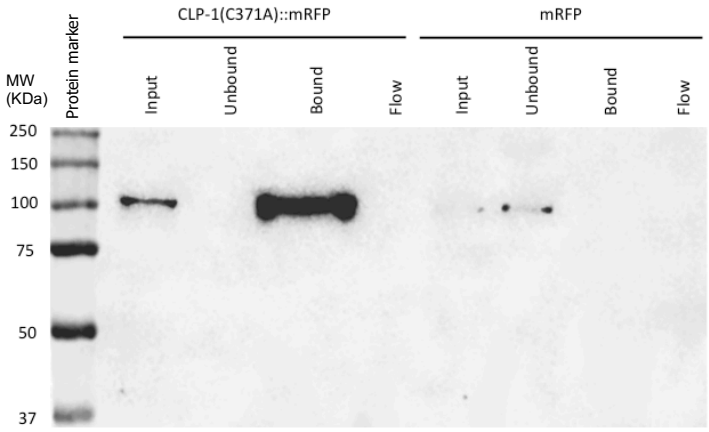


Figure 23. Beta-integrin co-immunoprecipitates with CLP-1(C371A)::mRFP.
Immunoblot of IP fractions of CLP-1(C371A)::mRFP and mRFP stained with anti-beta integrin antibody MH25. Beta-integrin (~110Kda) was enriched in the bound fraction of the immunoprecipitation of CLP-1(C371A)::mRFP but not of mRFP alone.

3.2 Determining a minimal localisation domain of CLP-1.

Overview

The *C. elegans* sarcomere (Figure 9) consists of actin thin filaments anchored to the cell membrane via attachment to dense bodies, which are homologous in function to vertebrate z-lines**.** Actin thin filaments are interspersed with myosin thick filaments, which are anchored to the cell membrane by M-lines. M-lines form rows that are inter-digited with dense bodies (74,75). It was previously shown that a CLP-1::mRFP fusion protein driven from its native promoter localises to sarcomeric M-lines and cell contact sites in body wall muscle (5).

To identify the region(s) of the CLP-1 protein that is responsible for promoting its localisation to m-lines and adhesion junctions, a series of plasmid clones expressing truncated CLP-1::mRFP fusion proteins under control of *unc-54* muscle specific promoter were constructed. *unc-54* encodes myosin heavy chain II, which is primarily expressed in body wall muscle cells (74). It was hypothesised that the series of truncated CLP-1 proteins would serve to delineate the region(s) responsible for localisation to sarcomere m-lines and adjacent to dense bodies. Truncated versions of CLP-1 were expressed as mRFP fusion proteins and the localisation pattern of each was examined by confocal microscopy. The design of truncations was guided by the domain architecture of CLP-1. All plasmids were constructed; however not all plasmids were microinjected.

3.2.1 Construction of mRFP tagged CLP-1 truncation expression plasmids

CLP-1 truncations were designed to delete singly or in combination the domains; glycine rich (GR) region and domain 1 (CLP-1Δ1), the catalytic CysPC domain (CLP-

1Δ2) and domain III (CLP-1Δ3) (Figure 24). Domain boundaries of CLP-1 were defined based on a multiple sequence alignment performed using Clustal Omega with other atypical *C. elegans* calpains, human typical calpains CAPN1, CAPN2, Drosophila CALPA (Figure 25) (127).To minimise structural disruption to the non-truncated domains, 25 amino acids adjacent to the boundary were included from the disrupted domain in question. The majority of these adjacent amino acids belong to domain linker regions with no predicted secondary structure based on alignment with the crystal structure of CAPN2 (Figure 25).

The plasmid clones pLN3, pLN4, pLN8, pLN9, pLN10 and pLN5 (Figure 24) were cloned using site directed PCR deletion mutagenesis starting with pMC7 as template. pMC7 encodes *clp-1* cDNA with 3' mRFP fusion under the control of muscle specific *unc-54* promoter.

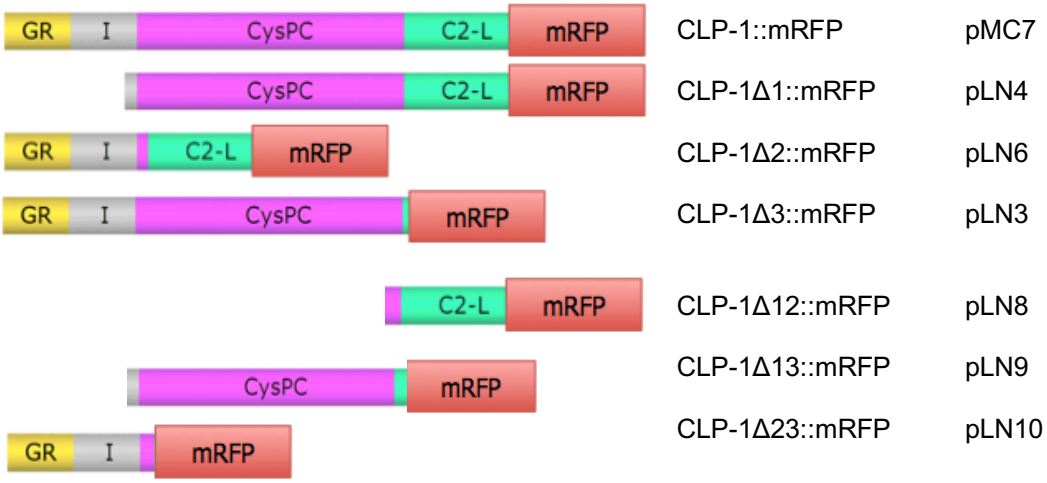
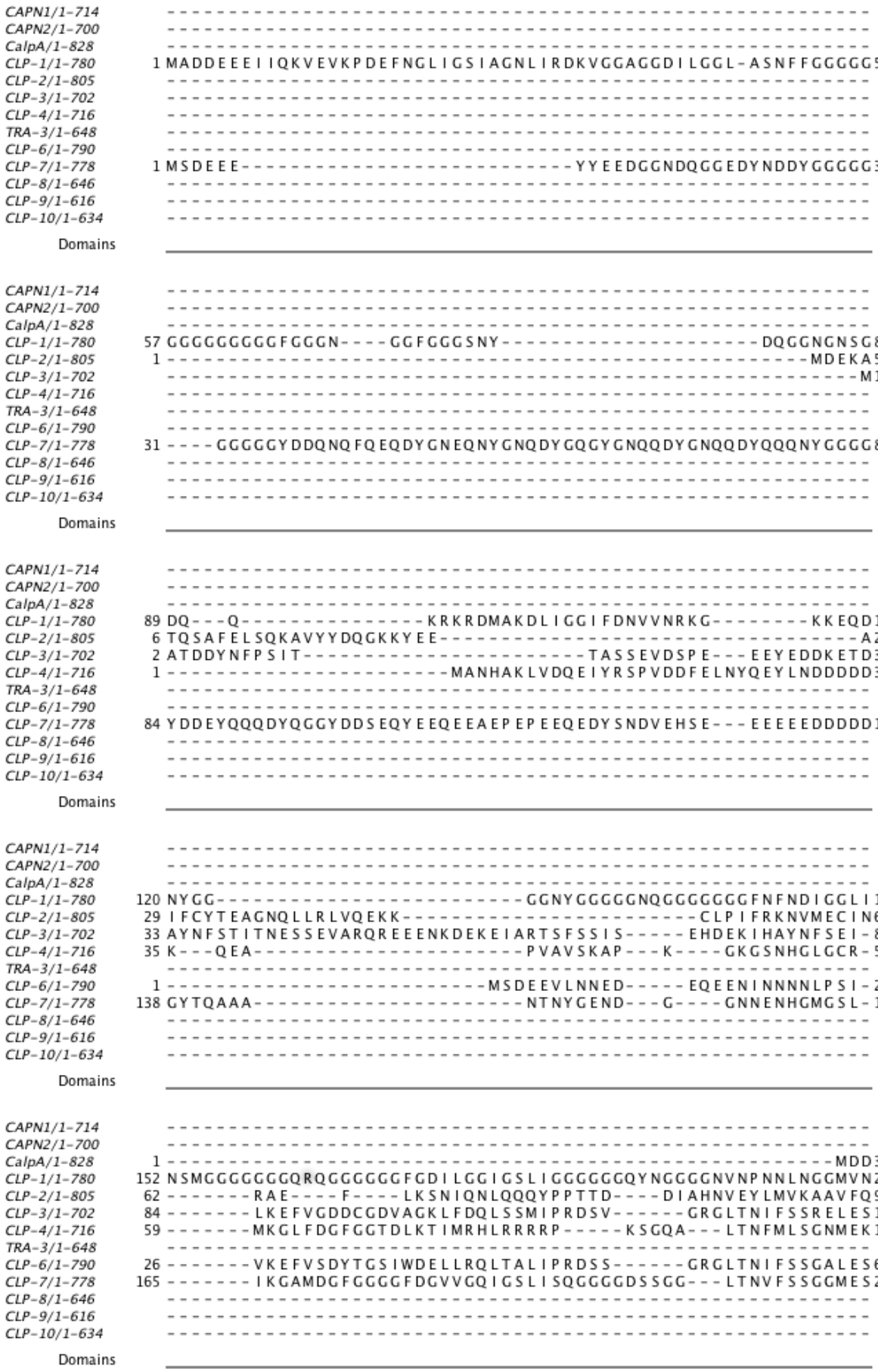
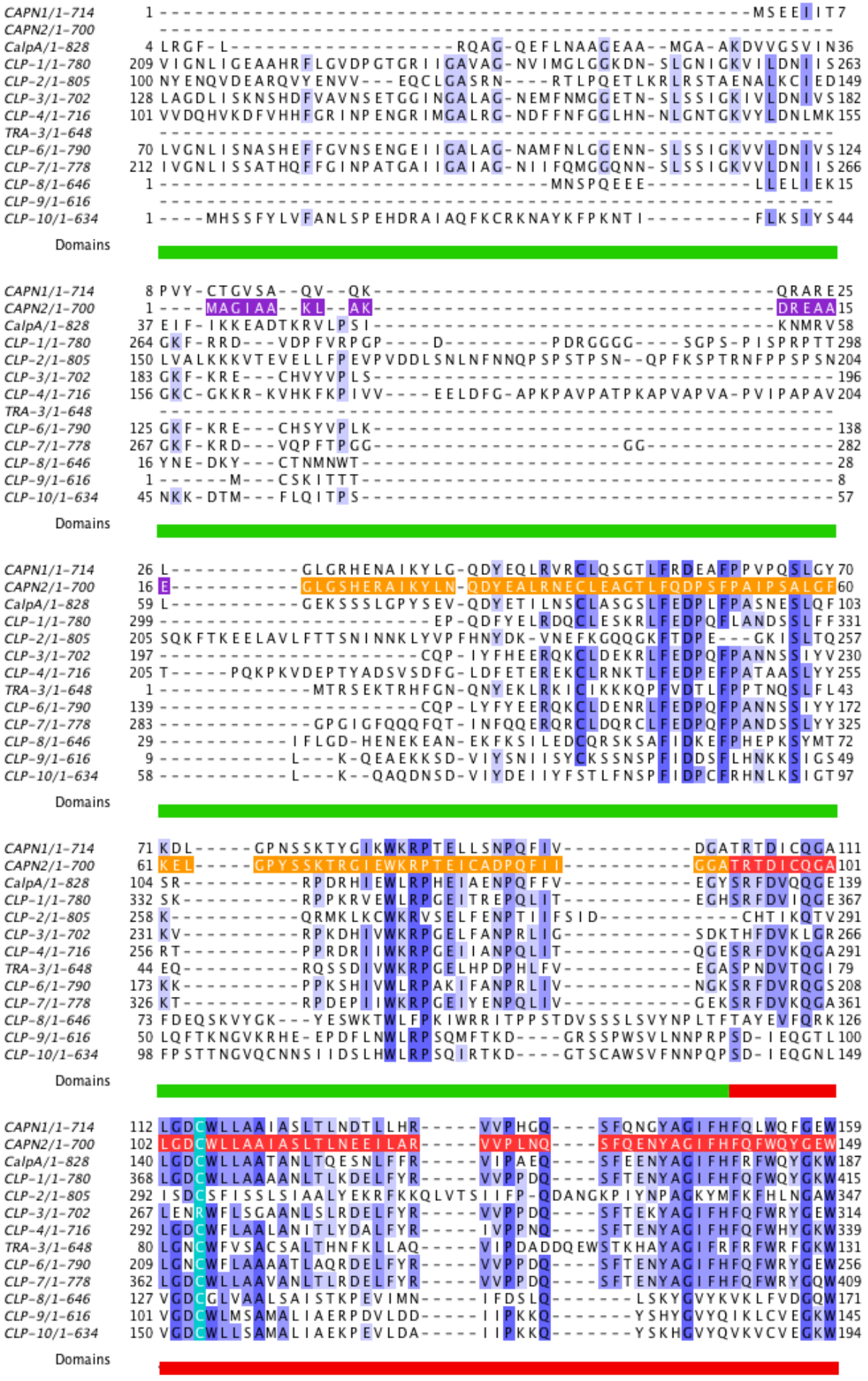


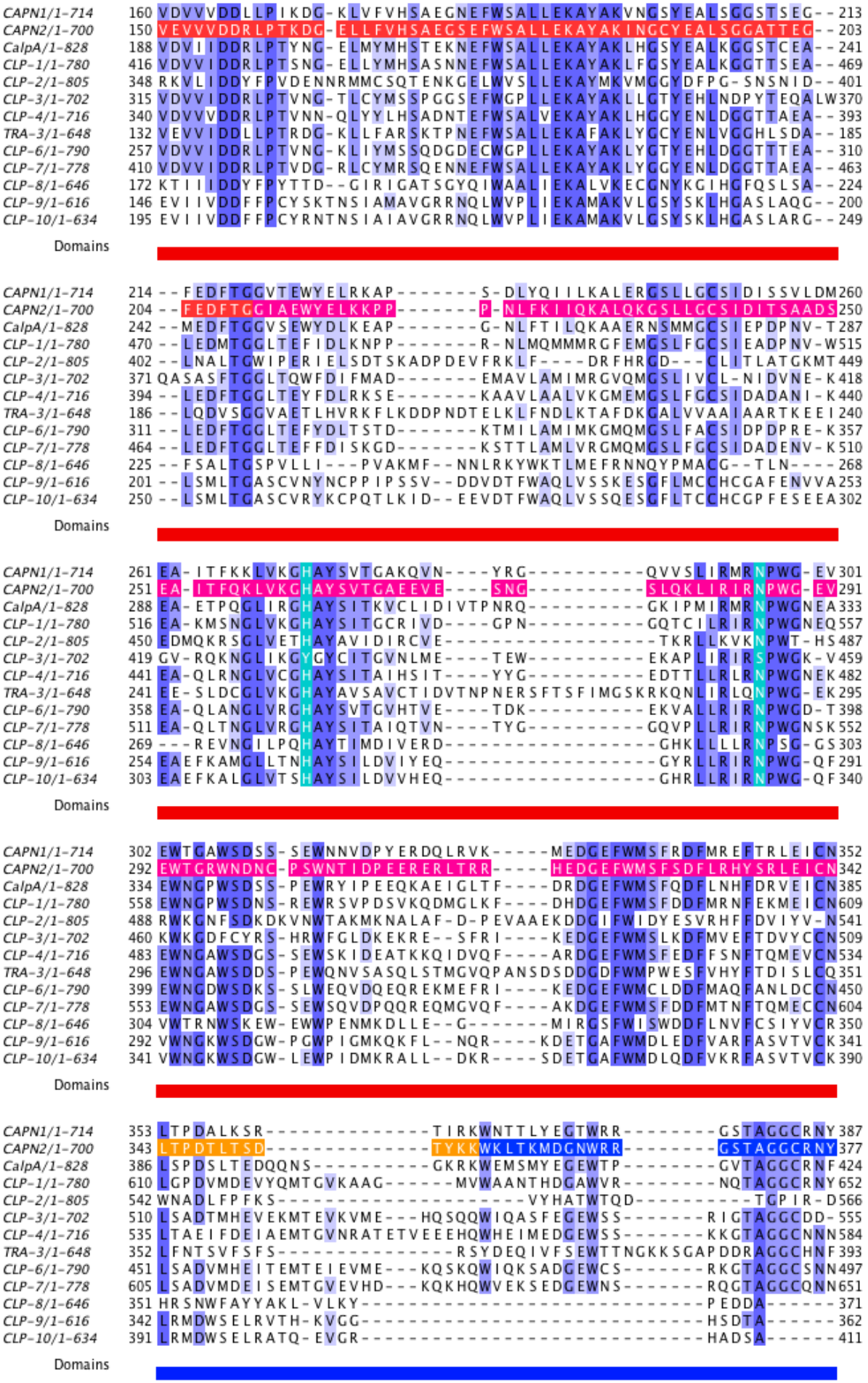
Figure 24. CLP-1 truncation mRFP fusion proteins.

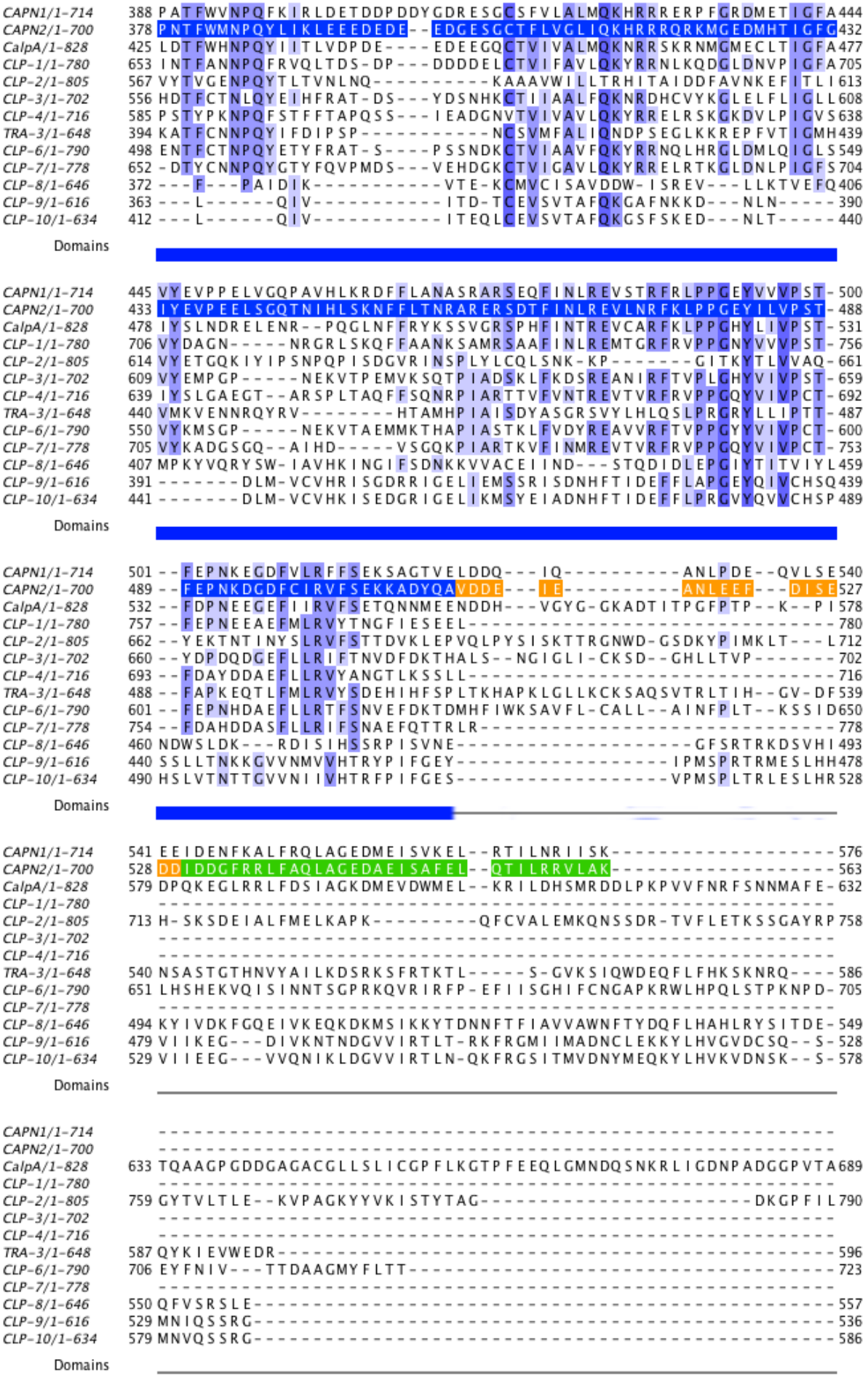
Domain maps of wild type CLP-1 fused at the C-terminus to mRFP and the truncated fusion proteins CLP-1Δ1::mRFP, CLP-1Δ2::mRFP, CLP-1Δ3::mRFP, CLP-1Δ12::mRFP , CLP-1Δ13::mRFP and CLP-1Δ23::mRFP designed to delete different regions of CLP-1.

Multiple sequence alignment of calpain proteins.




CAPN1/1-714
CAPN2/1-700
CalpA/1-828
CLP-1/1-780
CLP-2/1-805
CLP-3/1-702
CLP-4/1-716
TRA-3/1-648
CLP-6/1-790
CLP-7/1-778
CLP-8/1-646
CLP-9/1-616
CLP-10/1-634
Domains


CAPN1/1-714 160 VDVVVDDLLPIKDG-KLVFVHSAEGNEFWSALLEKAYAKVNGSYEALSGGSTSEG--213
CAPN2/1-700 150 VEVVVDDRLPTKDG-ELLFVHSAEGSEFWSALLEKAYAKINGCYEALSGGATTEG--203
CalpA/1-828 188 VDVIIDDRLPTYNG-ELMYMHSTEKNEFWSALLEKAYAKLHGSYEALKGGSTCEA--241
CLP-1/1-780 416 VDVVIDDRLPTSNG-ELLYMHSASNNEFWSALLEKAYAKLFGSYEALKGGTTSEA--469
CLP-2/1-805 348 RKVLIDDYFPVDENNRMMCSQTENKGELWVSLLEKAYMKVMGGYDFPG-SNSNID--401
CLP-3/1-702 315 VDVVIDDRLPTVNG-TLCYMSSPGGSEFWGPLLEKAYAKLLGTYEHLNDPYTEQALW370
CLP-4/1-716 340 VDVVVDDRLPTVNN-QLYYLHSADNTEFWSALVEKAYAKLHGGYENLDGGTTAEA--393
TRA-3/1-648 132 VEVVIDDLLPTRDG-KLLFARSKTPNEFWSALLEKAFAKLYGCYENLVGGHLSDA--185
CLP-6/1-790 257 VDVVIDDRLPTVNG-KLIYMSSQDGDECWGPLLEKAYAKLYGTYEHLDGGTTTEA--310
CLP-7/1-778 410 VDVVIDDRLPTVDG-RLCYMRSQENNEFWSALLEKAYAKLYGGYENLDGGTTAEA--463
CLP-8/1-646 172 KTIIIDDYFPYTTD--GIRIGATSGYQIWAALIEKALVKECGNYKGIHGFQSLSA--224
CLP-9/1-616 146 EVIIVDDFFPCYSKTNSIAMAVGRRNQLWVPLIEKAMAKVLGSYSKLHGASLAQG--200
CLP-10/1-634 195 EVIIVDDFFPCYRNTNSIAIAVGRRNQLWVPLIEKAMAKVLGSYSKLHGASLARG--249
Domains
CAPN1/1-714 214 --FEDFTGGVTEWYELRKAP-------S-DLYQIILKALERGSLLGCSIDISSVLDM260
CAPN2/1-700 204 --FEDFTGGIAEWYELKKPP-------P-NLFKIIQKALQKGSLLGCSIDITSAADS250
CalpA/1-828 242 --MEDFTGGVSEWYDLKEAP-------G-NLFTILQKAAERNSMMGCSIEPDPNV-T287
CLP-1/1-780 470 --LEDMTGGLTEFIDLKNPP-------R-NLMQMMMRGFEMGSLFGCSIEADPNV-W515
CLP-2/1-805 402 --LNALTGWIPERIELSDTSKADPDEVFRKLF----DRFHRGD---CLITLATGKMT449
CLP-3/1-702 371 QASASFTGGLTQWFDIFMAD-------EMAVLAMIMRGVQMGSLIVCL-NIDVNE-K418
CLP-4/1-716 394 --LEDFTGGLTEYFDLRKSE-------KAAVLAALVKGMEMGSLFGCSIDADANI-K440
TRA-3/1-648 186 --LQDVSGGVAETLHVRKFLKDDPNDTELKLFNDLKTAFDKGALVVAAIAARTKEEI240
CLP-6/1-790 311 --LEDFTGGLTEFYDLTSTD-------KTMILAMIMKGMQMGSLFACSIDPDPRE-K357
CLP-7/1-778 464 --LEDFTGGLTEFFDISKGD-------KSTTLAMLVRGMQMGSLFGCSIDADENV-K510
CLP-8/1-646 225 --FSALTGSPVLLI---PVAKMF--NNLRKYWKTLMEFRNNQYPMACG--TLN----268
CLP-9/1-616 201 --LSMLTGASCVNYNCPPIPSSV--DDVDTFWAQLVSSKESGFLMCCHCGAFENVVA253
CLP-10/1-634 250 --LSMLTGASCVRYKCPQTLKID--EEVDTFWAQLVSSQESGFLTCCHCGPFESEEA302
Domains
CAPN1/1-714 261 EA-ITFKKLVKGHAYSVTGAKQVN----YRG----------QVVSLIRMRNPWG-EV301
CAPN2/1-700 251 EA-ITFQKLVKGHAYSVTGAEEVE----SNG----------SLQKLIRIRNPWG-EV291
CalpA/1-828 288 EA-ETPQGLIRGHAYSITKVCLIDIVTPNRQ----------GKIPMIRMRNPWGNEA333
CLP-1/1-780 516 EA-KMSNGLVKGHAYSITGCRIVD----GPN----------GQTCILRIRNPWGNEQ557
CLP-2/1-805 450 EDMQKRSGLVETHAYAVIDIRCVE----------TKRLLKVKNPWT-HS487
CLP-3/1-702 419 GV-RQKNGLIKGYGYCITGVNLME----TEW----------EKAPLIRIRSPWGK-V459
CLP-4/1-716 441 EA-QLRNGLVCGHAYSITAIHSIT----YYG----------EDTTLLRLRNPWGNEK482
TRA-3/1-648 241 EE-SLDCGLVKGHAYAVSAVCTIDVTNPNERSFTSFIMGSKRKQNLIRLQNPWG-EK295
CLP-6/1-790 358 EA-QLANGLVRGHAYSVTGVHTVE----TDK----------EKVALLRIRNPWGD-T398
CLP-7/1-778 511 EA-QLTNGLVRGHAYSITAIQTVN----TYG----------GQVPLLRIRNPWGNSK552
CLP-8/1-646 269 ---REVNGILPQHAYTIMDIVERD----------GHKLLLLRNPSG-GS303
CLP-9/1-616 254 EAEFKAMGLLTNHAYSILDVIYEQ----------GYRLLRIRNPWG-QF291
CLP-10/1-634 303 EAEFKALGLVTSHAYSILDVVHEQ----------GHRLLRIRNPWG-QF340
Domains
CAPN1/1-714 302 EWTGAWSDSS-SEWNNVDPYERDQLRVK-----MEDGEFWMSFRDFMREFTRLEICN352
CAPN2/1-700 292 EWTGRWNDNC-PSWNTIDPEERERLTRR-----HEDGEFWMSFSDFLRHYSRLEICN342
CalpA/1-828 334 EWNGPWSDSS-PEWRYIPEEQKAEIGLTF----DRDGEFWMSFQDFLNHFDRVEICN385
CLP-1/1-780 558 EWNGPWSDNS-REWRSVPDSVKQDMGLKF----DHDGEFWMSFDDFMRNFEKMEICN609
CLP-2/1-805 488 RWKGNFSDKDKVNWTAKMKNALAF-D-PEVAAEKDDGIFWIDYESVRHFFDVIYV-N541
CLP-3/1-702 460 KWKGDFCYRS-HRWFGLDKEKRE--SFRI----KEDGEFWMSLKDFMVEFTDVYCCN509
CLP-4/1-716 483 EWNGAWSDGS-SEWSKIDEATKKQIDVQF----ARDGEFWMSFEDFFSNFTQMEVCN534
TRA-3/1-648 296 EWNGAWSDDS-PEWQNVSASQLSTMGVQPANSDSDDGDFWMPWESFVHYFTDISLCQ351
CLP-6/1-790 399 EWNGDWSDKS-SLWEQVDQEQREKMEFRI----KEDGEFWMCLDDFMAQFANLDCCN450
CLP-7/1-778 553 EWNGAWSDGS-SEWSQVDPQQREQMGVQF----AKDGEFWMSFDDFMTNFTQMECCN604
CLP-8/1-646 304 VWTRNWSKEW-EWWPENMKDLLE--G-------MIRGSFWISWDDFLNVFCSIYVCR350
CLP-9/1-616 292 VWNGKWSDGW-PGWPIGMKQKFL--NQR----KDETGAFWMDLEDFVARFASVTVCK341
CLP-10/1-634 341 VWNGKWSDGW-LEWPIDMKRALL--DKR----SDETGAFWMDLQDFVKRFASVTVCK390
Domains
CAPN1/1-714 353 LTPDALKSR----------TIRKWNTTLYEGTWRR--------GSTAGGCRNY387
CAPN2/1-700 343 LTPDTLTSD----------TYKKWKLTKMDGNWRR--------GSTAGGCRNY377
CalpA/1-828 386 LSPDSLTEDQQNS----------GKRKWEMSMYEGEWTP--------GVTAGGCRNF424
CLP-1/1-780 610 LGPDVMDEVYQMTGVKAAG------MVWAANTHDGAWVR--------NQTAGGCRNY652
CLP-2/1-805 542 WNADLFPFKS----------VYHATWTQD----------TGPIR-D566
CLP-3/1-702 510 LSADTMHEVEKMTEVKVME---HQSQQWIQASFEGEWSS-------RIGTAGGCDD-555
CLP-4/1-716 535 LTAEIFDEIAEMTGVNRATETVEEEHQWHEIMEDGEWSS-------KKGTAGGCNN584
TRA-3/1-648 352 LFNTSVFSFS----------RSYDEQIVFSEWTTNGKKSGAPDDRAGGCHNF393
CLP-6/1-790 451 LSADVMHEITEMTEIEVME---KQSKQWIQKSADGEWCS-------RKGTAGGCSNN497
CLP-7/1-778 605 LSADVMDEISEMTGVEVHD---KQKHQWVEKSEDGEWNS-------RQGTAGGCQNN651
CLP-8/1-646 351 HRSNWFAYYAKL-VLKY----------PEDDA------371
CLP-9/1-616 342 LRMDWSELRVTH-KVGG----------HSDTA------362
CLP-10/1-634 391 LRMDWSELRATQ-EVGR----------HADSA------411
Domains


CAPN1/1-714
CAPN2/1-700
CalpA/1-828
CLP-1/1-780
CLP-2/1-805
CLP-3/1-702
CLP-4/1-716
TRA-3/1-648
CLP-6/1-790
CLP-7/1-778
CLP-8/1-646
CLP-9/1-616
CLP-10/1-634
Domains

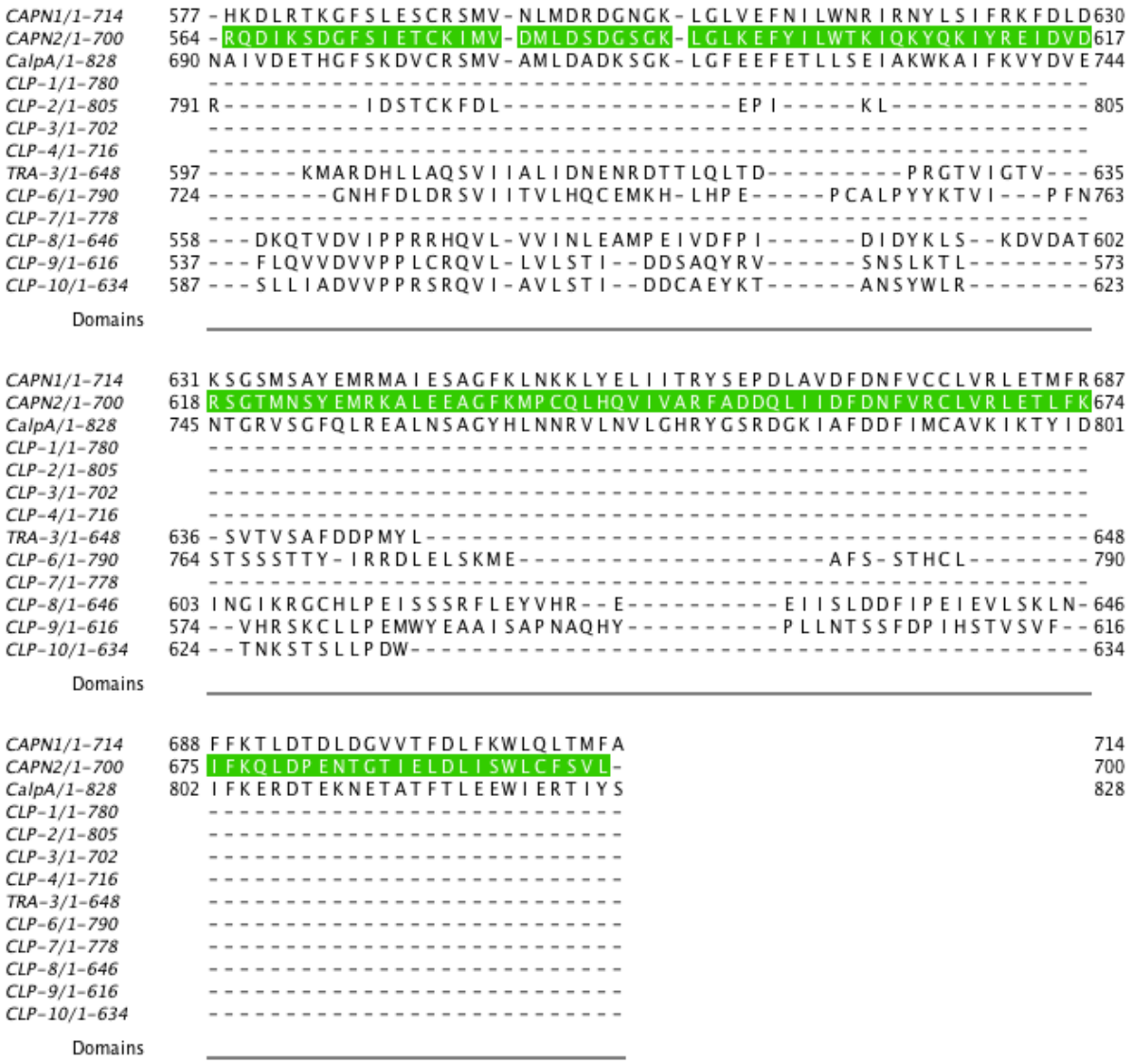


Figure 25. Multiple sequence alignment of calpain proteins.

Protein sequence alignment of *C. elegans* atypical calpains (CLP-1 to CLP-10) human calpains (CAPN1 and CAPN2) and *D. melanogaster* CALPA. Domain I (green) Domain II (red) Domain III (blue) are highlighted below the alignment. Conserved catalytic residues C, H and N are highlighted in cyan. CLP-3 lacks these catalytic residues. CAPN2's sequence is highlighted with domains and linker regions attributed from its crystal structure (128); Domain I (purple), DI->DII linker (orange), Domain IIa (red), Domain IIb (pink), DII->DIII linker (orange), DIII (blue), EF hand domain (green). Proteins were aligned using Clustal Omega and shaded using Jalview 2.82. Dark blue boxes highlight greater than 80% sequence similarity, blue boxes >60% sequence similarity light blue boxes >40% sequence similarity and uncoloured boxes less than 40% sequence similarity.

Plasmid clones pLN3, pLN4, pLN8, pLN10 and pLN5 were constructed similarly using one step site directed deletion mutagenesis with different primer pairs (Figures 26, 27) to express different truncations of *clp-1* cDNA fused to mRFP at the C-terminus and also mRFP alone, under the control of the *C. elegans unc-54* muscle specific promoter. pLN9, a plasmid that only expresses domain II was constructed by ligating two PCR products; the expression vector was amplified using primer pair PK1317/PK1323 from pMC7 and the domain II containing insert was amplified using primer pair PK1810/PK1811 from template plasmid pMC7 (Figure 28).

The site directed deletion mutagenesis PCR products were then re-circularised and verified by using bacterial colony PCR or restriction digest. Minipreps of pLN3, pLN4 and pLN5 were identified by Sma1 restriction digest (Figure 29). Bacterial colony PCR using primers PK1356 and PK1355 that span the entire *clp-1* cDNA region were used to verify the successful truncation of domains from pLN6, pLN8, pLN9 and pLN10 (Figures 30, 31). For all clones, the regions spanning the deletions were sequenced to confirm in frame truncation of the domain in question. The plasmids pLN3, pLN4, pLN5, pMC7 were microinjected into the germline of *C. elegans* PK2969 hermaphrodites (Method 2.4.2). The plasmids cloned and strains produced from microinjection are summarized below in Table 11.

Plasmid	Description	Cloned	Strains
pLN3	Expresses CLP-1Δ3::mRFP that has a truncated domain 3 of CLP-1.	Yes	PK3144
pLN4	Expresses CLP-1Δ1::mRFP that has a truncated N-terminal glycine rich region and domain 1 of CLP-1.	Yes	PK3228
pLN6	Expresses CLP-1Δ2::mRFP that has a truncated domain 2 of CLP-1.	Yes	N/A
pLN8	Expresses CLP-1Δ12::mRFP that has truncated N-terminal glycine rich region, domains 1 and 2 of CLP-1.	Yes	N/A
pLN9	Expresses CLP-1Δ12::mRFP that has truncated N-terminal glycine rich region, domains 1 and 3 of CLP-1.	Yes	N/A
pLN10	Expresses CLP-1Δ12::mRFP that has truncated domains 2 and 3 of CLP-1.	Yes	N/A

Table 11. Summary of the CLP-1 truncation clones that were generated and injected to produce transmitting lines.
Strain PK3144 displayed fluorescence and was imaged by confocal microscopy, PK3228 exhibited no fluorescence.

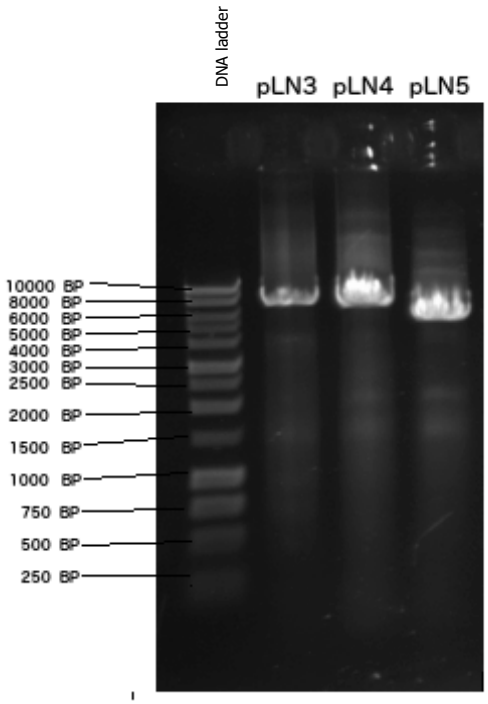


Figure 26. PCR products: pLN3, pLN4 and pLN5

PCR products amplified from pMC7; pLN3 is a 7419 bp PCR product with a disrupted CLP-1 domain 1. pLN4 is a 6906 bp PCR product disrupted domain 3. pLN5 is a 5502 bp PCR product that has no CLP-1 (*unc-54::mRFP*:: *unc-54* 3' UTR*)*.

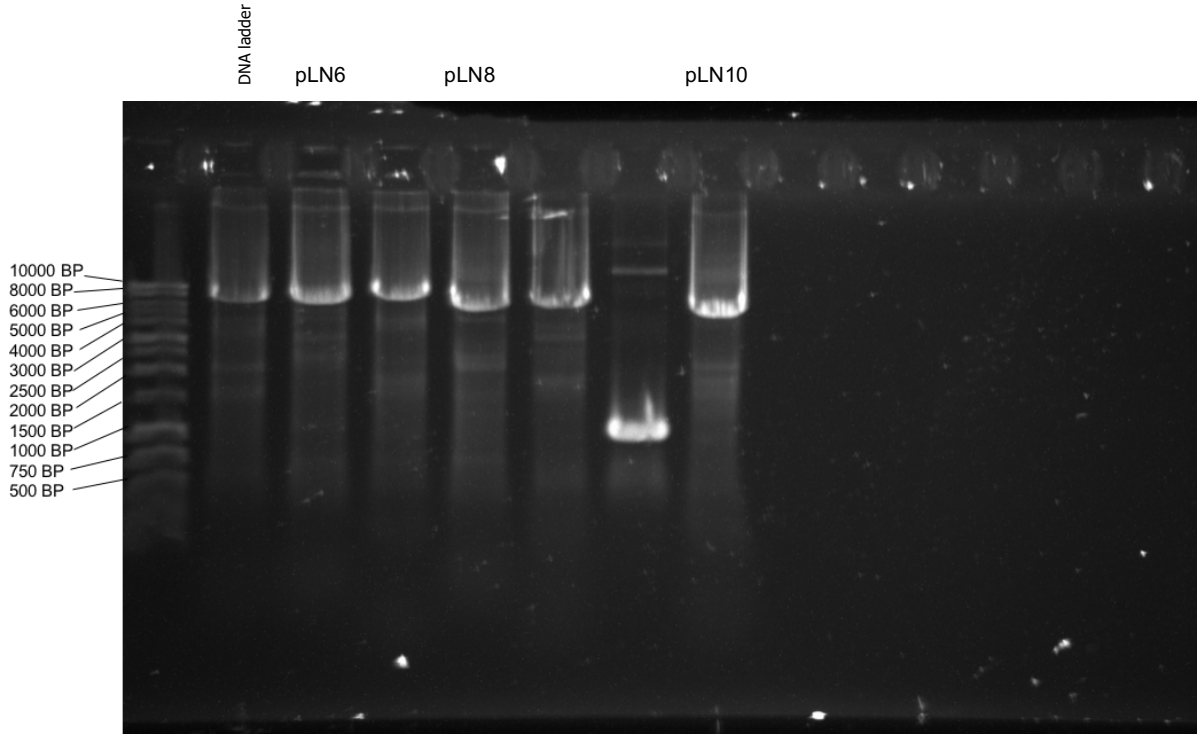


Figure 27. pLN6, pLN8 and pLN10 linear PCR products generated by site directed deletion mutations of pMC7.

pLN6 is a 6984 bp product using primer pair PK1667 and PK1666 in a PCR on pMC7 template causing a truncation of domain II. pLN8 is a 6021 bp product using primer pair PK1323 and PK1666 causing truncation of domains I and II. pLN10 is a 6468 bp product using primer pair PK1667 and PK1317 causing truncation of domains II and III.

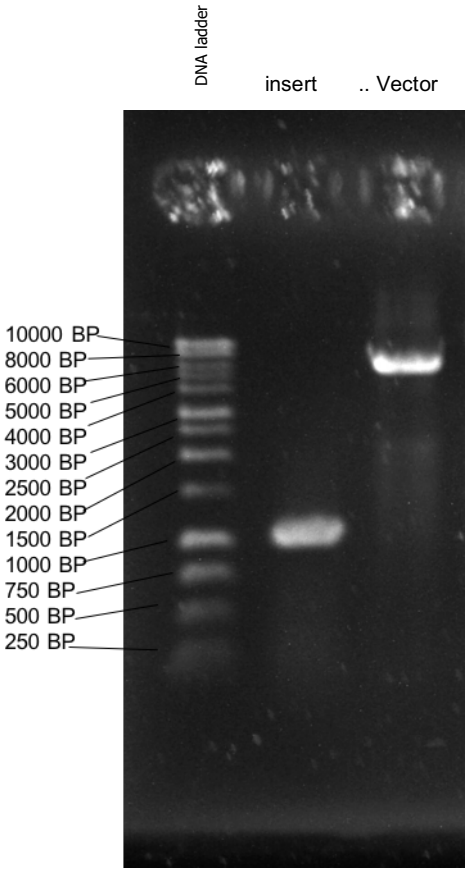


Figure 28. pLN9 PCR products used for Gibson assembly cloning.

1015bp Insert containing domain II was amplified from pMC7 using primers using PK1810 and PK1811. 5505bp expression vector containing *unc-54* promoter was amplified using PK1317 and PK1323.

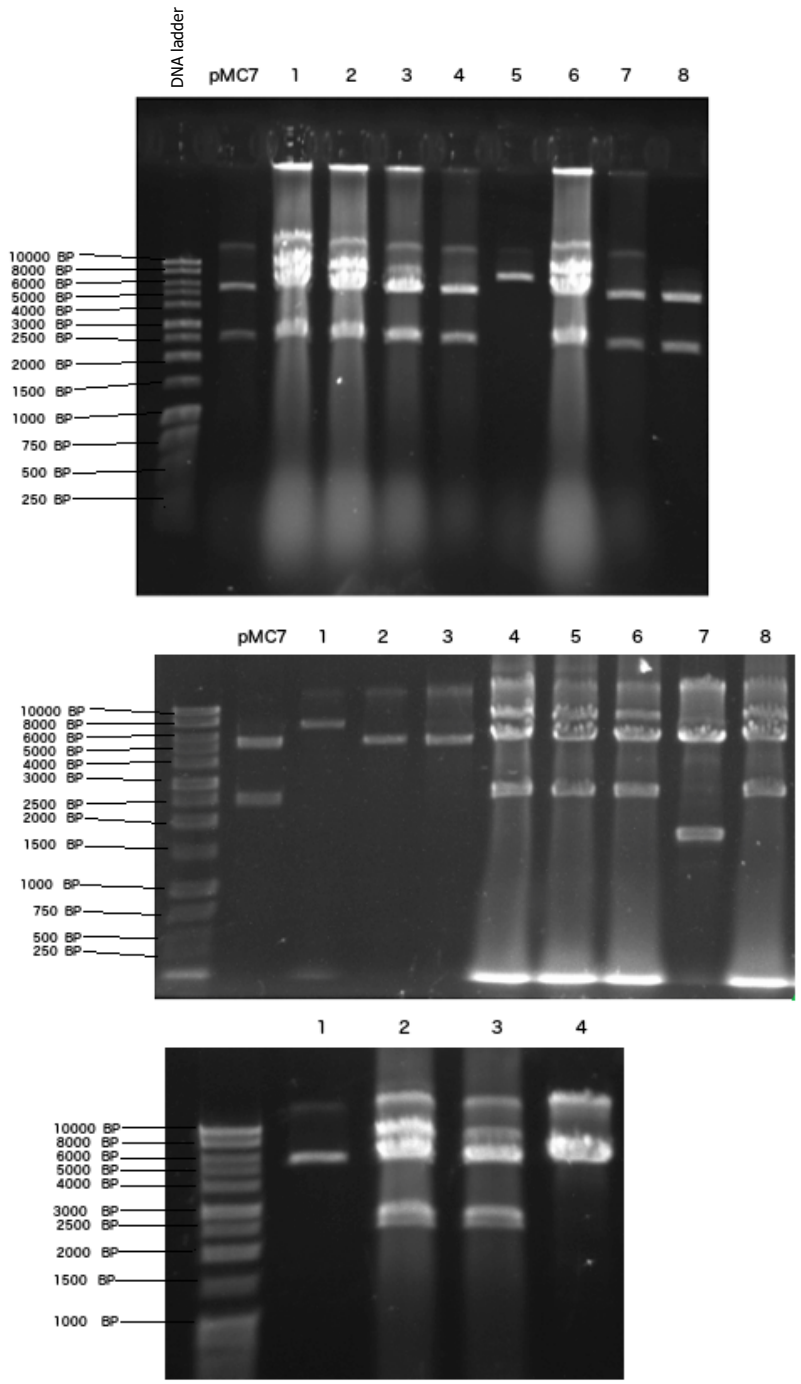


Figure 29. Sma1 digests of pLN3, pLN4 and pLN5.

A. Sma1 digest of plasmid DNA isolated from *E. coli* XL-1 blue colonies transformed with pLN3, '5' has the correct band of 7419 bp. The bands from other colony minipreps correspond to the PCR template plasmid pMC7. B. Sma1 digest of plasmid DNA isolated from *E. coli* XL-1 blue colonies transformed with pLN4, '7' has the predicted band pattern of a band of 5374 bp and another of 1532 bp corresponding to pLN4. C. Sma1 digest of plasmid DNA isolated from XL-1 blue colonies transformed with pLN5, '1' and '4' have the correct single band of 5502 bp.

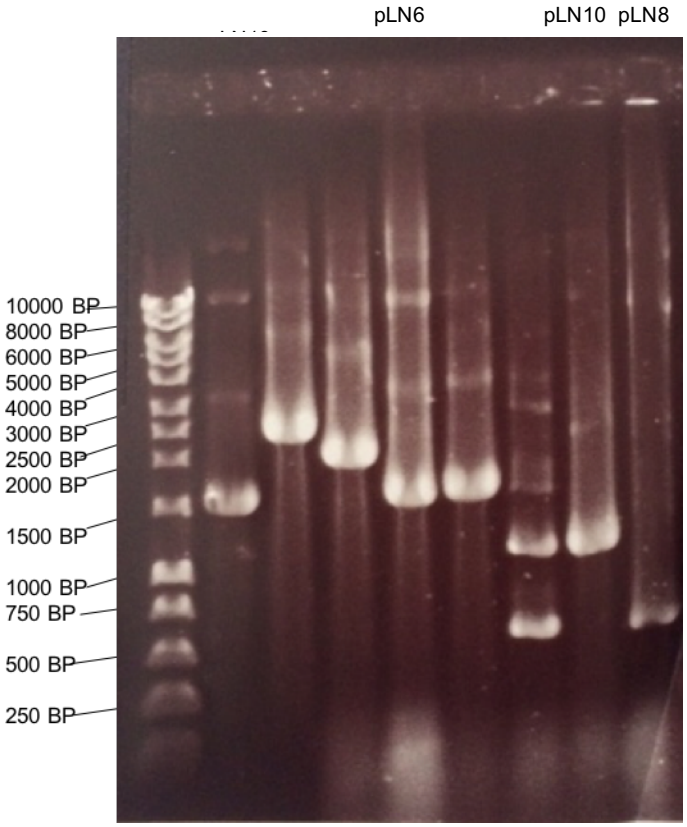


Figure 30. Bacterial colony PCR confirming domain truncations of plasmids pLN6, plL8, pLN10 using primer pair PK1355 and PK1356.

The amplicon produced by PCR using primer pair PK1355 and PK1356 spans the entire *clp-1* cDNA. PCR using PK1355 and PK1356 on clones pLN6, pLN8 and pLN10 was used to confirm presence of truncation prior to confirmation by sequencing. Using PK1355 and PK1356 for PCR; truncation of domain II in pLN6 results in a 1692 bp product, truncation of domain I and II results in a 666bp product and truncation of pLN10 results in a 1113 bp product.

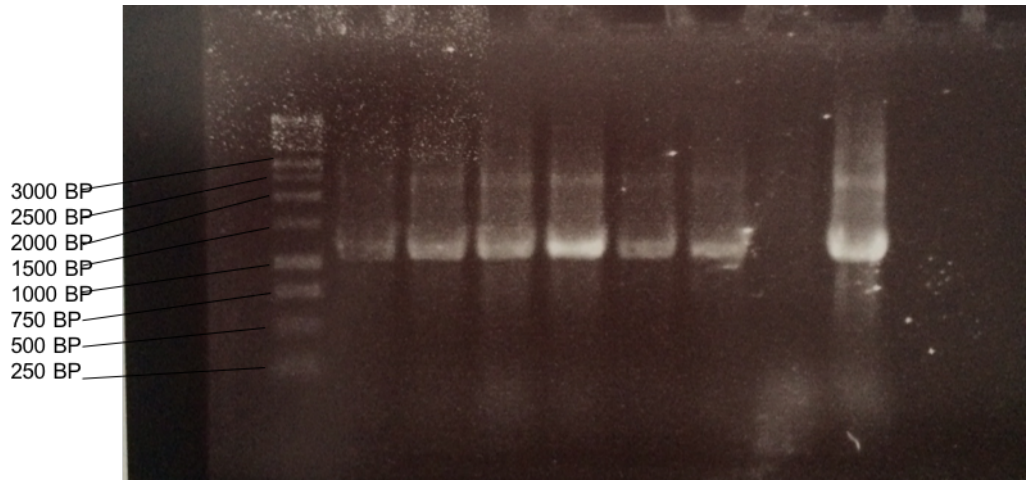


Figure 31. Bacterial colony PCR of pLN9 clones using primers PK1356 and PK1355.

The amplicon produced by PCR using primer pair PK1355 and PK1356 spans *clp-1* cDNA, truncation of domains I and III produces a 1125 bp product using primers PK1356 and PK1355 in a PCR.

PK2969 expresses a PAT-3::GFP reporter, which is associated with dense bodies, m-lines and adhesion plaques, enabling easier identification of intracellular structures. From the microinjections stable lines transmitting extrachromosomal arrays were obtained; PK3144 [*unc-54::mRFP crEX491*], PK3146 [*unc-54::clp-1Δ3::mRFP crEx499*], PK3228 [*unc-54::clp-1Δ1::mRFP crEx501*], Pk3083 [*unc-54::clp-1::mRFP crEX468*]. The presence of target plasmids within extrachromosomal arrays were confirmed by single worm PCR.

3.2.2 Active CLP-1 and inactive CLP-1 localise to m-lines and adjacent to dense bodies in the sarcomere

Confocal laser scanning microscopy (CLSM) was used to visualise the intracellular expression of fluorescent proteins in body wall muscle. The intracellular localisation of strain PK3114 [*unc-54::clp-1(C371A)::mRFP crIs18 X*] expressing catalytically inactive CLP-1 was also analysed.

Although pLN4 was successfully injected to generate strain PK3228, PK3228 exhibited no fluorescence. It is possible that this lack of fluorescence was due to a mutation in pLN4 in a region outside of the region spanning the deletion that was verified by sequencing. There was insufficient time to inject the other clones. Images of the muscle cells captured using a laser scanning confocal microscope (LSCM) showed that CLP-1::mRFP, CLP-1(C371A)::mRFP and CLP-1 were localised diffusely throughout the cytoplasm while being excluded from the nucleus (Figure 32). pPJ83 was also injected and strain PK3118 [*clp-1p::clp-1::GFP* + *pCoel::mRFP crEX484*] expressing CLP-1::GFP under its native promoter was imaged (Figure 32) and similar diffuse cytoplasmic localisation was observed. GFP only expressed under a muscle specific promoter was also imaged (strain BC10095).

As shown in Figure 33 panels A-F, mRFP by itself was diffusely expressed at the base of the sarcomere but like CLP-1::mRFP it was excluded from the dense bodies. CLP-1::mRFP, CLP-1(C371A)::mRFP and CLP-1Δ3::mRFP were enriched in m-lines and areas immediately adjacent to dense bodies (Figure 33, panels D-L).

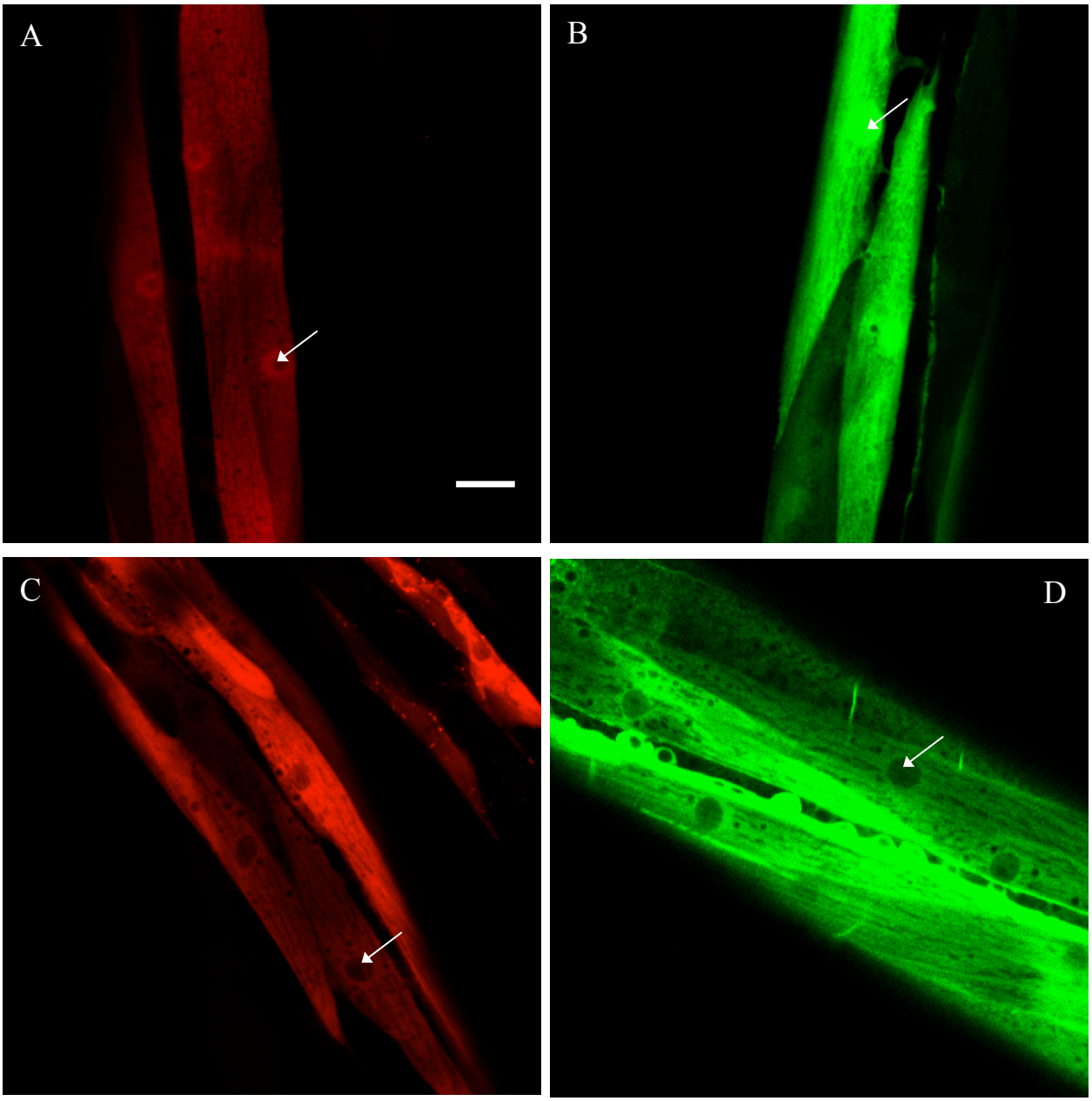


Figure 32. Intracellular expression pattern of CLP-1 fluorescent reporters.

A. mRFP and B. GFP expressed under control of muscle specific promoters are present throughout the cytoplasm and also accumulate in nuclei of body wall muscle cells. C . CLP-1::mRFP and D.CLP-1::GFP are present in the cytoplasm of body wall muscle cells but are excluded from the nucleus. Mitotic loss of extrachromosomal transgenes causes the mosaic pattern of expression observed in B and C (86). Arrowheads indicate body wall muscle cell nuclei. Scale bar is 10 μm.

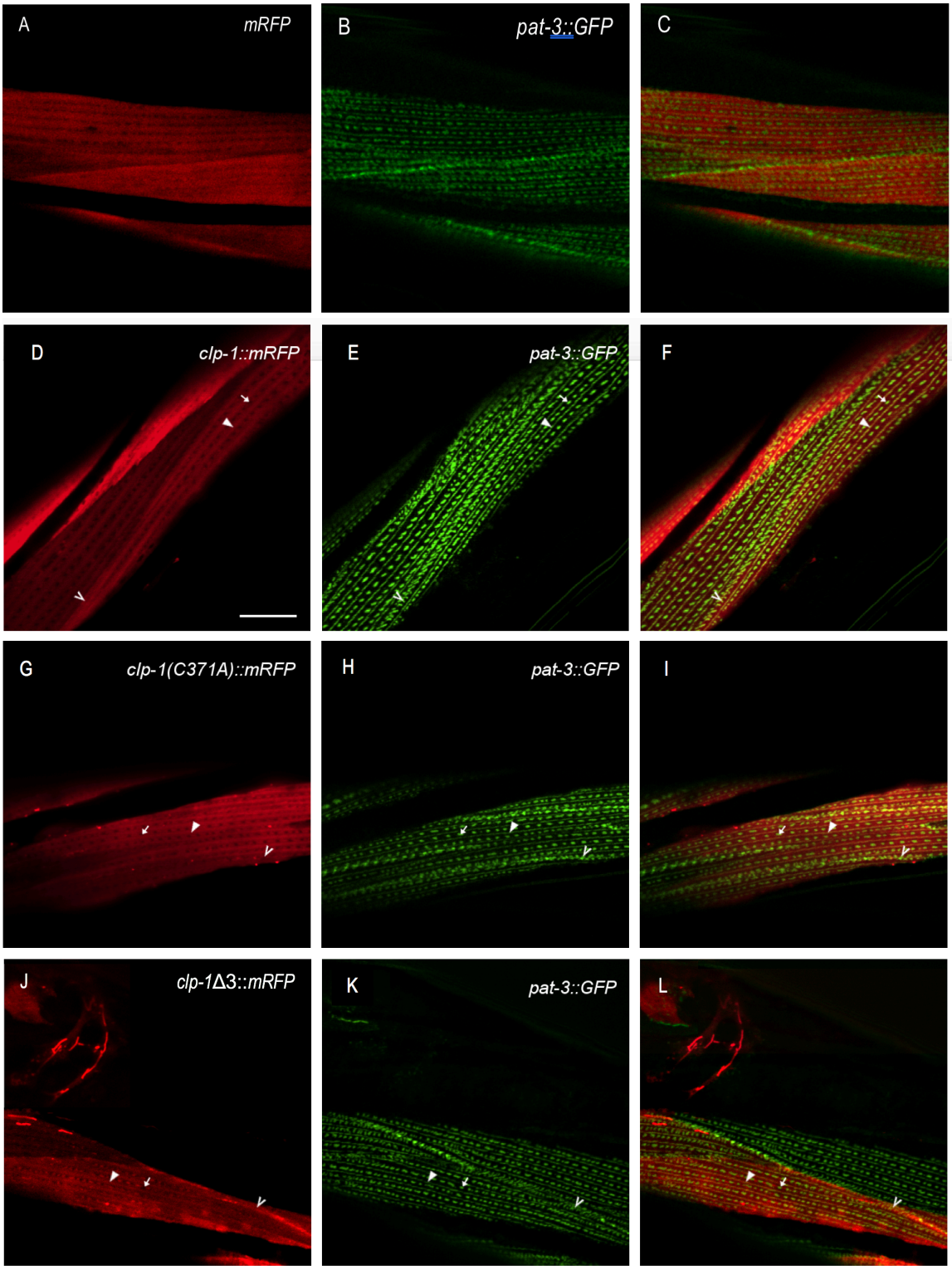


Figure 33. Intracellular localisation of CLP-1.

CLP-1 localises to sarcomere m-lines and adhesion plaques in body wall muscle. (A) CLP-1::mRFP, (D) CLP-1(C371A)::mRFP, (G) CLP-1Δ3::mRFP (J) mRFP are expressed in body wall muscle. (B, E, H, K) PAT-3::GFP localises to myofilament-sarcolemma attachment complexes; m-lines, dense bodies and adhesion plaques. (C) mRFP is excluded from the dense bodies. (F) CLP-1::mRFP, (I) CLP-1(C371A)::mRFP, (L) CLP-1Δ3::mRFP colocalise with PAT-3::GFP at m-lines and

adhesion plaques and are enriched in between adjacent dense bodies, but not localised to dense bodies. Dense body, small arrow; arrowhead, M-lines; open arrowhead, adhesion plaques. Scale bars are 10 μM.

3.3 Physiological analysis of calpains: Involvement of calpains in heat stress and attempted construction of a *clp-1/clp-4* double knockout mutant.

Overview:

Heat induced stress in *C. elegans* was shown to induce an increase of calcium ion concentration in the muscle of *C. elegans* and adversely affect the locomotion and survival of adult worms (95). It was shown that adult *C. elegans* subjected to heat stress of 2-3 hours at 35°C had significant increase of intracellular Ca^{2+} levels and a decrease in locomotion and survival rates (129). Raising the temperature of the heat stress to 39°C led to widespread necrotic cell death (130). Resistance to heat stress including the suppression of increased calcium ion levels was observed in *daf-2* long-life mutants which have activated DAF-16 (129).

Heat stress causes a similar phenotype to overexpression of CLP-1 in body wall muscle and increases calcium ion levels. Thus, it might be expected that high temperature would increase the activity of calpain due to the rise in Ca^{2+} levels. It was investigated whether calpain induced paralysis might be increased in heat stress conditions. To determine whether calpains are involved in the heat-stress pathogenesis the mutant CLP-1 overexpression transgenic PK3163 [*unc-54::clp-1::mRFP crIs19*] strain, CLP-1 deletion mutant PK2341 [*clp-1(tm690)*] strain and wildtype animals were subjected to heat stress for 2 hours at 35°C. The animals were then examined for paralysis/death over the subsequent 3 days post heat stress and comparison of the percentage of paralysis/death exhibited by each strain was made. The percentage of paralysed/dead worms was also recorded for each strain after 3 days at 20°C for comparison.

3.3.1 Calpain is an aggravating factor in heat induced stress.

Paralysed and dead worms were recorded at t=0 hours prior to heat stress, t=2 hours after heat stress and every 24 hours (from t=0) for 3 days. At t=2 hours, after heat stress, it was observed that all worms were paralysed. In agreement with the observations of Momma et al. (95), it was observed that 26.7% (n = 60) WT adult worms exhibited paralysis/death after 3 days post heat stress at 35°C (Table 12). This percentage increased to 55% in heat stressed PK3163 worms overexpressing CLP-1::mRFP (Figure 34), indicating that CLP-1 overexpression increases heat stress induced paralysis/death (P-value <0.05).

For non-heat stressed PK3163 worms overexpressing CLP-1::mRFP, 21.7% paralysed/dead worms were observed after 3 days (Table 13) and this increased to 55% in heat stressed PK3163 worms (Figure 34). This result suggests that paralysis caused by calpain overexpression is increased (P-value <0.05, Fisher's exact test) in heat stressed worms. Momma et al. (95) found that heat stress induced a spike in intracellular calcium ion levels. This might cause an increase in calpain catalytic activity. It is also possible that other pathological pathways are activated/upregulated during heat stress in combination with calpain over-activation.

The number of paralysed/dead *clp-1(tm690)* worms did not significantly differ from wild type worms after 3 days under normal conditions (20°C) or 3 days after heat stress (Table 12, Table 13, Figure 34).

There was insufficient time to repeat this experiment multiple times.

genotype	WT movement	paralysed	dead
	0 hours		
WT	60	0	0
clp-1(tm690)	60	0	0
[*unc-54::clp-1::mRFP*]	60	0	0
	2 hours		
WT	0	60	0
clp-1(tm690)	0	60	0
[*unc-54::clp-1::mRFP*]	0	60	0
	1 day		
WT	59	0	1
clp-1(tm690)	54	1	5
[*unc-54::clp-1::mRFP*]	49	4	7
	2 days		
WT	56	0	4
clp-1(tm690)	54	0	6
[*unc-54::clp-1::mRFP*]	36	10	14
	3 days		
WT	44	9	7
clp-1(tm690)	45	3	12
[*unc-54::clp-1::mRFP*]	27	17	16

Table 12. Paralysis and survival of worms subjected to heat stress of 35°C for 2 hours.
An increased number of worms were paralysed/died after 3 days ($P < 0.05$, Fisher's exact test) expressing the CLP-1::mRFP transgene compared to wild type worms after 2 hours of heat stress. The number of paralysed/dead *clp-1(tm690)* worms did not significantly differ from wild type worms after 3 days, following heat stress.

genotype	WT movement	paralysed	dead
	0 hours		
WT	60	0	0
clp-1(tm690)	60	0	0
[*unc-54::clp-1::mRFP*]	60	0	0
	2 hours		
WT	60	0	0
clp-1(tm690)	60	0	0
[*unc-54::clp-1::mRFP*]	60	0	0
	1 day		
WT	60	0	0
clp-1(tm690)	60	0	0
[*unc-54::clp-1::mRFP*]	51	5	4
	2 days		
WT	60	0	0
clp-1(tm690)	60	0	0
[*unc-54::clp-1::mRFP*]	48	3	6
	3 days		
WT	59	1	0
clp-1(tm690)	59	1	0
[*unc-54::clp-1::mRFP*]	47	5	8

Table 13. Paralysis and survival of worms maintained at 20°C.
A greater number of worms were paralysed/died after 3 days expressing the CLP-1::mRFP transgene compared to wild type worms maintained at 20°C. The number of paralysed/dead *clp-1(tm690)* worms did not significantly differ from wild type worms after 3 days at 20°C.

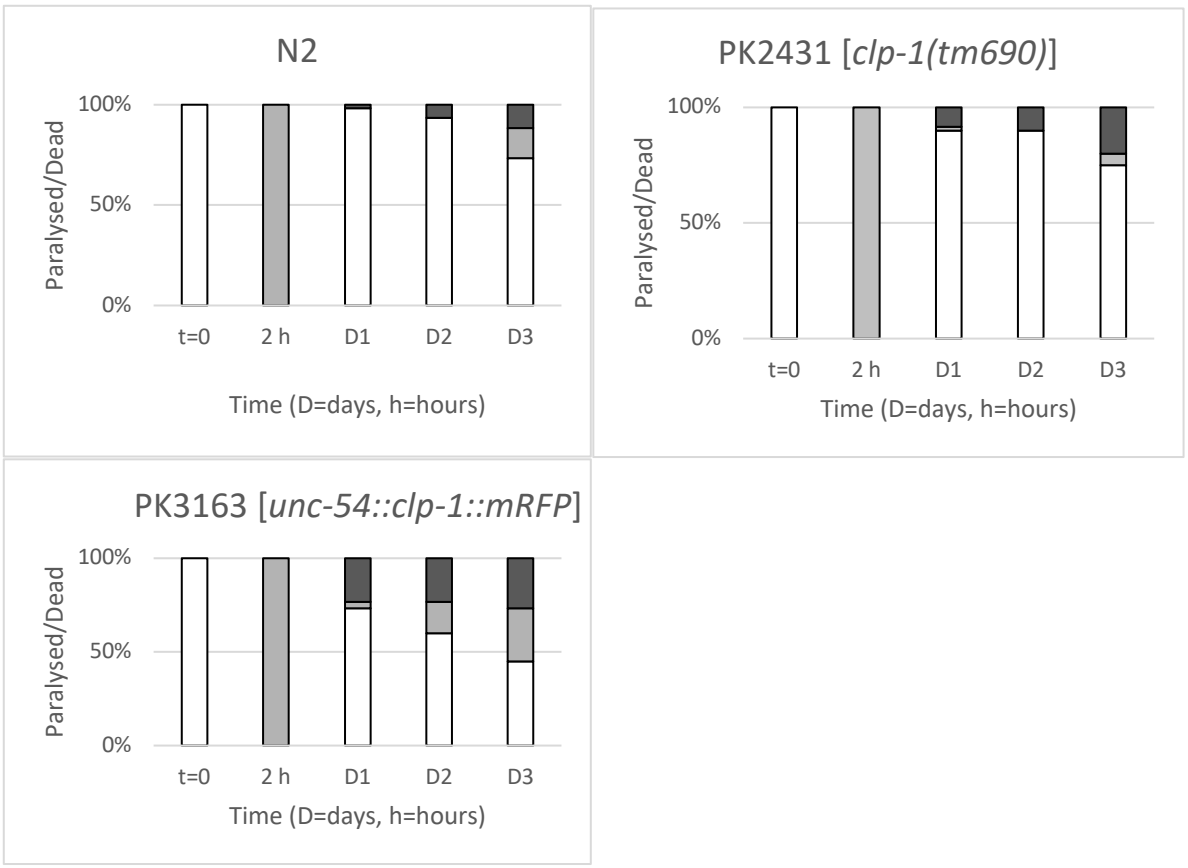


Figure 34. Paralysis and survival rates of heat stressed wildtype (N2), *clp-1(tm690)* and *unc-54::clp-1::mRFP* worms.

White: Wild type mobility. Light grey: paralysis. Dark grey: dead. All worms were paralysed after the initial 2 hour heat stress treatment at 35°C. The percentage of paralysed/dead worms (n=60) was greater for adult PK3163 worms vs N2 worms. The percentage of paralysed/dead) PK2431 worms did not significantly differ from wild type worms after 3 days at 20°C.

3.3.2 Attempted construction of a *clp-1/clp-4* double knockout mutant by CRISPR-Cas9.

Previously RNAi knockdown of muscle expressed calpains *clp-1* or *clp-4* showed no obvious phenotype in adult *C. elegans*. It was hypothesized this might be the consequence of functional redundancy (5). Construction of a *clp-1;clp-4* double mutant using the CRISPR-Cas9 system was attempted but not completed.

To create a *clp-1;clp-4* double mutant, CRISPR-Cas9 genetic engineering was attempted to create a 374 bp deletion of *clp-4* was attempted in *clp-1* deletion mutant strain PK2341 [*clp-1(tm690)*] (Figure 35). PK2341 has a 624 bp deletion mutation in *clp-1*, and is predicted to express a protein lacking a complete catalytic triad and catalytic activity. Protospacer adjacent motif (PAM) sites were selected with aid of an online CRISPR-Cas9 design tool (131) and an online primer maker tool that I created named CrisprPrimerMaker (Figure 36). CrisprPrimerMaker automates a stage of sgRNA primer design for CRISPR-Cas9, following the method of Arribere et al. (132). PAM sites were targeted either side of the region to be deleted, and plasmids expressing single guide RNAs (sgRNA) targeting these sites were cloned (Figures 37).

28 worms were injected and 8 of these had progeny that contained dumpy or roller worms indicative of CRISPR-Cas9 activity. 50 worms were screened from one of the broods exhibiting the greatest number of dumpy and roller worms (42 dumpy and 3 roller) by PCR, part of this screen is shown in Figure 38. None of the worms screened displayed a 309 bp band expected from deletion of the site targeted by the sgRNA.

It is possible that no mutants were generated due to inefficiency of the targeted sites at being cleaved by Cas9, however it might also be due to the rarity of successful CRISPR-Cas9 cut event and lack of screening/injection limited by time.

Antisense *clp-4* genomic:
tttatatcaaattttaagcaaattttaagcaatattatatcttagccactaaattaaaatgagaaatatgtcaaaataggcc
aaaaatctcaatttacCTGCATAATTCTCCGTAAAACTTTGATT/TGGCGGCACAATTCG
ATAAAACAACGCATCATACAACGTGATATTTGCCAGAGCCGCCAAAAACCAGCA
ATCACCCAATGCTCCCTGCTTCACATCAAATCTCGATTCTCCTTGCGTAATCAGC
TGTGGATTGGCAATAATCTCACCGGGGCGTTTCCAAATTATTCTGTCACGTGGT
GGGGTACGGTAGTAAAGGGATGCGGCGGTTGCTGGAAACTCTGGATCCTCAAA
TAGTGTCTTATTCCGGAGACATTTCTCACGTTCTGTCTCGAAATCTAGGCCGAA
ATCGGAGACGGAATCGGCGTAGGTTGGCTCGTCAActgaaaattctggaaattggaattcag
cgtgacgagagctttcaaataacacccatcatg/acgaggttacgaaaatt

lowercase : introns.
UPPERCASE: EXONS.

Yellow highlight: protospacer (length 19nt)
Blue highlight: PAM site
Purple highlight: Cas9 cut site
Red highlight: cysteine of catalytic triad

Figure 35. Protospacer adjacent motifs (PAM) sites targeted for *clp-4* deletion mutation.

Two PAM sequences were chosen to target the catalytic triad of *clp-4* resulting in a deletion of a 374 bp region resulting in a frameshift mutation.

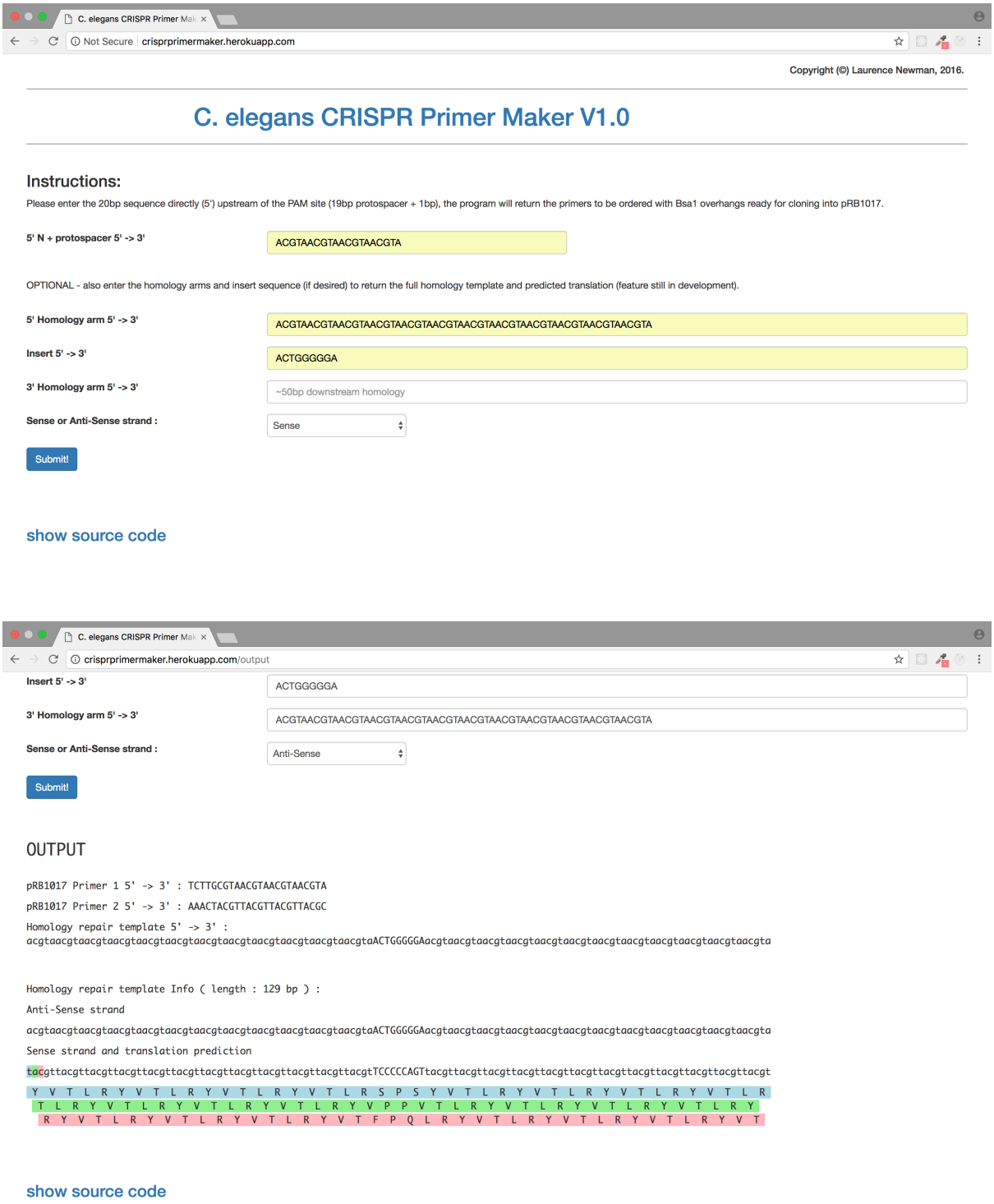


Figure 36. Web interface and output of CrisprPrimerMaker. Automates a part of CRISPR-Cas9 primer design for use in pRB1017 plasmids.

This tool automates a stage of sgRNA primer design for CRISPR-cas9 (following method of (132)) and prints back what mutation will be generated by the inputted primers and the predicted translation of the mutated gene http://crisprprimermaker.herokuapp.com/ (code repository available at https://github.com/lauri3new/CRISPR-primer-maker).

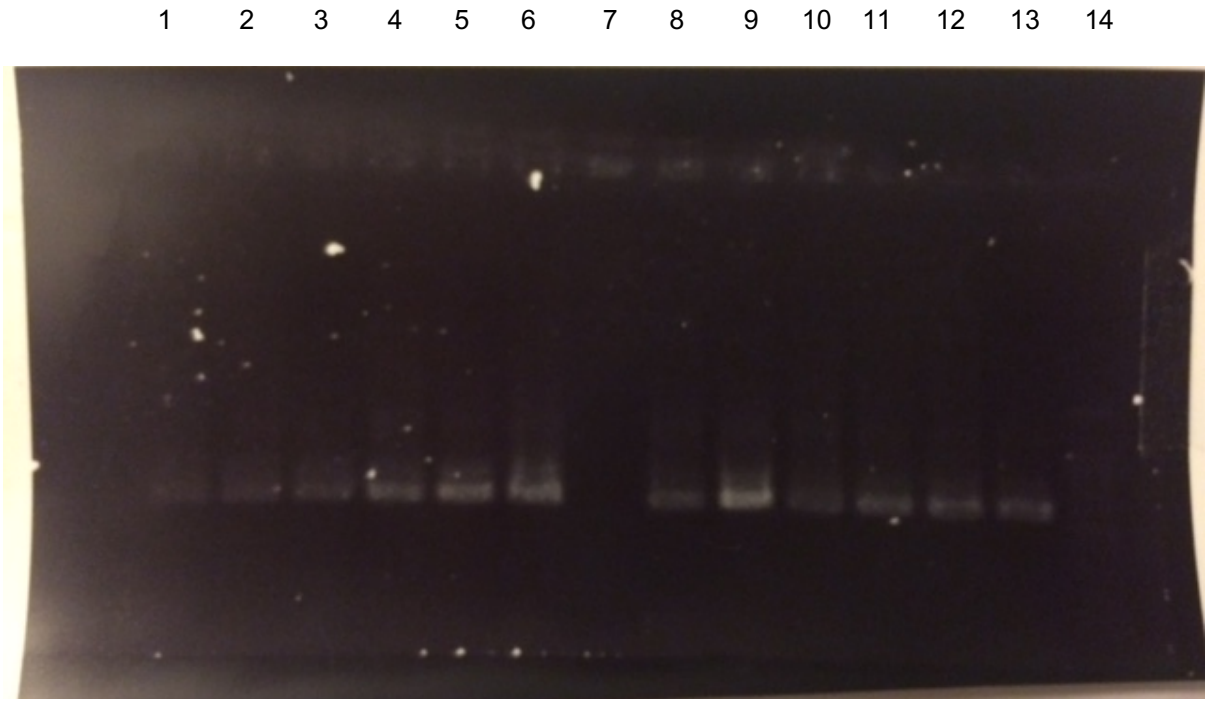


Figure 37. cloning of pRB1017 single guide RNA expression plasmids.

Bsa1 digested pRB1017 was ligated independently with annealed primers PK1813 and PK1858 (sgRNA plasmid 1) and with primers PK 1815 and PK1859 (sgRNA plasmid 2) to produce two sgRNA expression clones. Clones were verified by bacterial colony PCR using M13_forward primer and the gRNA insert antisense primer to produce 600 bp products (plasmid 1 lanes 2-6 and plasmid 2 lanes 8-13). Lane 14: DNA ladder.

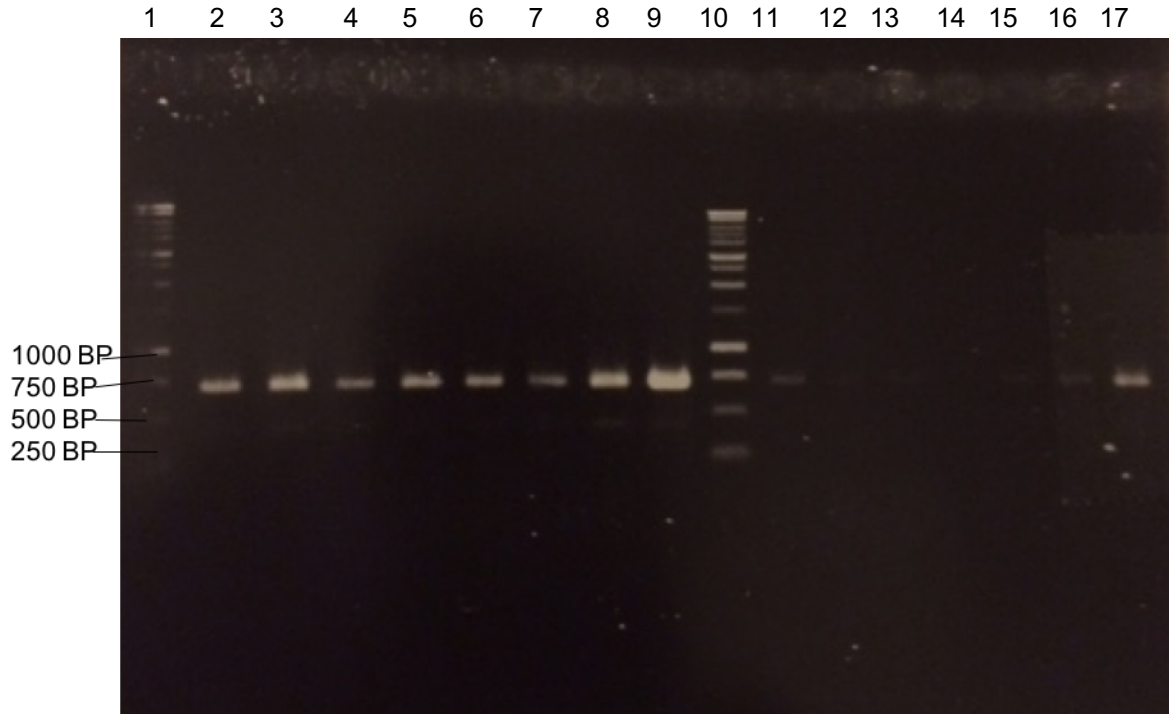


Figure 38. PCR screen for *clp-4* deletion mutation.

Bands at ~700bp indicate wild type *clp-4* gene, mutants would expect to have an extra band at ~300bp indicating deletion had occurred in one of the clp-4 genes. Bands 1 and 10 are DNA ladder. Bands 2-9, 11-17 are single worm PCR products indicating unsuccessful CRISPR-Cas9. No bands were observed in the 50 worms screened.

4 Conclusion and Discussion

4.1 Proteomic investigation into substrates of CLP-1

An integrated transgenic mutant strain PK3113 of *C. elegans* expressing CLP-1::(C371A)::mRFP and a control strain PK3114 expressing mRFP alone were generated. PK3113 and PK3114 integrated transgenes were mapped to the chromosomes II and X, respectively. Co-immunoprecipitation experiments using these strains were performed, and potential CLP-1 substrates were identified by differential comparisons.

Statistical analysis using Students t-test, SAINTexpress to calculate PSM fold change, and cut-off values identified 57 proteins that have the potential to be CLP-1 substrates. Potential substrates including orthologues of typical calpain substrates PCYT-1 and LMN-1, as well as novel interacting partners UNC-96 and TMEM-53 were amongst the 57 proteins enriched in the CLP-1(C371A)::mRFP co-IP compared to the control co-IP of mRFP alone. Calpain mediated degradation of these proteins might be responsible for muscle degradation and paralysis observed when CLP-1 is overexpressed in muscle. It would be expected that above normal cleavage of UNC-96 and PAT-3 by overactive CLP-1 in muscle cells would cause significant disruption to sarcomeres and release of myofibrillar proteins for degradation. If these proteins were subsequently validated as CLP-1 substrates this might in part account for the movement defects.

It is possible that false positives were identified amongst the 57 enriched proteins. For example, the proteasome regulatory protein RPN-13 is expressed in the nervous system, intestine and seam cells and not found in muscle cells (79). RPN-13 and other identified proteins from the proteomic screen might be artefacts caused during the lysis stage prior to IP, where CLP-1(C371A)::mRFP might have come in to contact with proteins from tissues other than muscle. However, CLP-1 is also expressed in other tissues including neurons under its endogenous promoter (5), so proteins expressed in tissues other than the muscle detected by the proteomic screen such as RPN-13 might still be identified from valid interactions and therefore potential substrates. False positives from the screen might also arise from other sources: the nature of the transgenic strain, the worm sample preparation as well as variation in the mass spectrometer used to capture MS spectra and identify proteins from the immunoprecipitations (133). There might have been false interactions caused by the increased transgenic expression of CLP-1::(C371A)::mRFP (96). It is also possible that the catalytically inactive CLP-1(C371A) has altered potential to bind substrates compared to wildtype CLP-1, therefore it is important that further validation studies are used to verify proteins identified from the proteomic screen as CLP-1 substrates. Given that CRISPR-Cas9 is now a viable option for generating transgenic worms, it would be of value to repeat the screen using CRISPR-Cas9 integrated CLP-1::mRFP. Single copy expression would likely more closely reproduce the endogenous amounts of CLP-1 found in muscle cells, and might lead to fewer false positives being identified in the co-IP MS analysis. However, the co-IP MS might also be less effective at identifying true positives as there would be lower concentration of CLP-1(C371A)::mRFP available for the RFP nanotrap (ChromoTek) to bind.

Previous work by Brule et al. (2) successfully identified typical calpain substrates from rat muscle tissue by immunoprecipitation of calpain using an anti-calpain antibody, followed by mass-spectrometry. The potential substrates identified from the calpain co-IP MS were then tested for co-localisation with CLP-1 in situ and in vitro cleavage by calpain.

Other techniques for substrate screening include ORF display and yeast 2 hybrid screens. In ORF display methods, a cDNA library expressing proteins with C-terminal biotin tags are bound to streptavidin and proteolysed by calpain causing release from the streptavidin. After sequential rounds of this process to eliminate non-specific targets substrates are identified by mass spectrometry (134). Yeast 2-hybrid assays were used to discover protein-protein interactions between calpains and proteins expressed from cDNA libraries (135). However, both these techniques are likely to yield false positives due to exposure of the calpain to proteins that might not usually co-localise in vivo. By using transgenic CLP-1::mRFP it is possible to focus the proteomic study by expressing under a tissue specific promoter. In this study CLP-1::mRFP was expressed under a muscle specific *unc-54* promoter. Another advantage of using transgenically expressed tagged CLP-1::mRFP is that the same worms can be used for imaging analysis. Immunoprecipitation of CLP-1(C371A)::mRFP was performed using an RFP nanotrap (ChromoTek). The RFP nanotrap (ChromoTek) has advantages over traditional antibodies used for co-immunoprecipitation experiments: it is derived from camelids and has a single light chain epitope recognising domain and it does not require protein A conjugation reducing the presence of contaminating proteins after immunoprecipitation of RFP

containing proteins. This is beneficial as fewer contaminating proteins are able to interfere with sample protein identification during mass spectrometry analysis.

From the set of proteins that were determined to be enriched in the CLP-1(C371A)::mRFP co-IP relative to the mRFP alone co-IP the most overrepresented GO terms included muscle system morphology variant. This is consistent with the hypothesis that CLP-1 is digesting muscle specific proteins that are important for structural and functional integrity leading to paralysis in the CLP-1 overexpression mutant. Under normal conditions CLP-1 might maintain proper muscle morphology by regulating the turnover of these proteins (10,77).

In addition to proteomic analysis, 3 candidate proteins were tested as potential CLP-1 substrates; intermediate filament, vinculin and *pat-3*. These proteins were tested for co-immunoprecipitation with CLP-1. PAT-3 co-immunoprecipitated with CLP-1 under high detergent conditions; however, it did not appear as an enriched protein in the proteomic screen. It is likely that it was not solubilised in the lower detergent lysis buffer that was used to preserve protein-protein interaction for the co-IP followed by LC-MSMS. The cytoplasmic domains of beta-integrin were shown to be proteolysed by the typical calpain CAPN2/CAPNS in vitro (101). Calpains are proposed to have a physiological role in cell motility by cleaving focal adhesion proteins including integrin (31). *pat-3*(RNAi) was shown in adult worms results to disrupt sarcomeres and animal movement (77).

The WormGRAB tool was produced that might be of use to other researchers when analysing AP-MS experiments in *C. elegans*. WormGRAB collates experimental

information about a protein dataset and assists in literature searches and could be used alongside other bioinformatics tools. It is available and free to access using all modern web browsers.

4.1.1 Proteomic investigation into substrates of CLP-1: future work

To continue this investigation substrate validation experiments could be performed by co-expressing GFP tagged proteins identified from the proteomic screen in: 1) active CLP-1 overexpressing mutants, 2) WT worms, 3) *clp-1* null mutants and 4) worms treated with calpain inhibitors. GFP tags are commonly used in the worm community for tagging genes for localisation studies and mutant strains are often available. After co-expressing potential substrates and CLP-1, degradation of the co-expressed target proteins could be determined by western blot analysis. It would be expected that CLP-1 substrates would be degraded in CLP-1 overexpressing mutants compared to a *clp-1* knockout mutant, or possibly in the presence of a non-specific calpain inhibitor. Cleavage products could be analysed by mass spectrometry to identify cleavage sites.

It would also be of interest to perform the immunoprecipitation followed by mass spectrometry under different lysis conditions in case important CLP-1 substrates were not detected from the screen as they were not soluble in the H100 lysis buffer used. For example, using a lysis buffer with higher salt and or detergent concentration might result in a different population of proteins, such as was observed with the solubility and co-immunoprecipitation of PAT-3.

It would be beneficial for heterologous expression CLP-1 in *E. coli* to yield a functional protease, as this would allow faster and simpler testing of candidate substrates by co-expressing epitope tagged candidate proteins with CLP-1 in the absence or presence of calcium. Previously Joyce (2008) attempted to express his-tagged CLP-1 using an IPTG inducible expression vector, however this yielded no soluble protein (136). Attempts to refold CLP-1 from inclusion bodies were also made, but yielded no functional CLP-1 proteins. I began another attempt at heterologous expression of CLP-1 in *E. coli* by cloning the plasmid pLN7 which expresses CLP-1 in an arabinose inducible expression vector (pBAD_myc_his_C), however I did not have time to perform any induction expression experiments. Other heterologous expression systems could be considered for co-expression of CLP-1 and potential substrates identified from the proteomic screen, Sokol and Kuwabara (2000) expressed *C. elegans* CLP-5 (TRA-3) and it's substrate TRA-2A in Sf9 cells using a baculovirus system (78).

An RNAi screen followed by phenotypic scoring of either locomotion or sarcomere integrity by microscopy (of the targets identified by the proteomic screen) would help to identify proteins required for muscle development/maintenance. It is possible that a substrate of CLP-1 might produce a functional product upregulating muscle degeneration and paralysis. This kind of relationship is similar to the mammalian calpain substrate autophagy factor ATG-5, which when cleaved by calpain results in a functional degradation product that upregulates apoptotic cell death (46). To investigate if any of the proteins had this relationship, an RNAi screen of proteins identified in the proteomic screen and subsequent phenotype scoring could be performed in a CLP-1 overexpression background. A suppression of the paralysis

phenotype observed in CLP-1 overexpressing worms would show that the protein in question likely produces a functional product that upregulates muscle degeneration.

There are many online bioinformatics tools used to analyse proteomic data. WormMine is a datastore of *C. elegans* that can be queried against to retrieve specified information about a protein dataset (137). To construct queries for WormMine it is useful to have SQL knowledge, however it also has a query builder interface that can be used to generate queries. Textpresso is a text mining tool with features that include co-reference searching at sentence level and searching by biological concepts (138). WormGRAB was created as an opinionated tool to obtain experimentally useful data from multiple sources without requiring complex configuration, simply using a web browser. The WormGRAB web application could be expanded and enhanced by using natural language processing for the PubMed search to determine relationships between bait and prey. It could also be expanded to allow user configuration on what data sources are used and the amount of data retrieved. It could also be expanded to cover proteins in other species as well as co-IP MS experiments that have multiple bait proteins.

4.2. Determining a minimal localisation domain of CLP-1

The plasmids pLN1, pLN3, pLN4, pLN5, pLN6, pLN8, pLN9 and pLN10 encoding CLP-1::mRFP with different domain truncations (and mRFP alone) under control of a muscle specific *unc-54* promoter were successfully cloned. Domain truncations were guided by a MSA of *clp-1* and other calpain genes.

By microinjection lines stably expressing and transmitting the transgenes for plasmids pLN1, pLN5, pLN3 were established however there was insufficient time to inject the other plasmids.

LSCM images were captured of the intracellular localisation of adult *C. elegans* body wall muscle sarcomeres for strains expressing CLP-1::mRFP, CLP-1(C371A)::mRFP, CLP-1Δ3::mRFP and mRFP. In agreement with what was previously reported by Joyce et al. (5) on the localisation pattern of CLP-1::mRFP, CLP-1 is cytoplasmic diffuse but is also enriched at sarcomere m-lines and regions adjacent to dense bodies. I found that CLP-1(C371A)::mRFP localises similarly to CLP-1::mRFP and is also enriched at sarcomere m-lines and regions adjacent to dense bodies suggesting that catalytic activity is not required for normal localisation (Figure 16). CLP-1Δ3::mRFP localises similarly to CLP-1::mRFP suggesting that Domain III is not required for this localisation.

A study characterising the intracellular expression patterns of 227 body wall muscle C-terminally tagged GFP proteins, not including CLP-1, discerned more than 15 unique patterns of localisation in body wall muscle cells (139). CLP-1 is characterised as a cytoplasmic protein, however confocal microscopy allows greater resolution of z-field (depth) and CLP-1::mRFP was also found to be enriched at structures immediately adjacent to dense bodies and excluded from dense bodies at the base of the sarcomere (Figure 16) (5).

4.2.1 Determining a minimal localisation domain of CLP-1: future work

This study could be continued by generating strains expressing all *clp-1* truncation constructs. It might also be of interest to image worms under heat stress or in high Ca^{2+} background such as in an *egl-19(gf)* mutant, and see if the distribution of CLP-1 changes within the sarcomere. Given that in vitro experiments observing expressed tagged truncated versions of CAPN1 found the C2-L Domain III was necessary for localisation to the plasma membrane under ionomycin induced activating conditions (140), it would be worthwhile to investigate if CLP-1 also undergoes translocation to the plasma membrane when intracellular Ca^{2+} levels are increased, and whether the C2-L domain III is also necessary for this process.

To gain understanding of the functional importance of the domains of CLP-1, future work could determine which domains are required for CLP-1 proteolytic activity. The activity of truncated CLP-1 proteins could be measured if a substrate is successfully identified, and if degradation of this substrate can be observed by western blot. This would enable comparative analysis of a substrates degradation pattern between worms expressing full length and truncated CLP-1::mRFP and would indicate if the truncated domain is required for catalytic activity, depending on whether the degraded bands were present or not (141).

4.3 Physiological analysis of calpains: Involvement of calpains in heat stress and attempted construction of a *clp-1/clp-4* double knockout mutant.

Momma et al. (95) found that intracellular Ca^{2+} levels are increased after heat stress. It is possible that calpains might be over activated during heat stress, and contribute to the paralysis and death exhibited by heat stressed worms (95). Enhanced paralysis/death was observed in worms exposed to heat stress that are overexpressing CLP-1 in body wall muscle compared to wild type animals. This might indicate that the level of active calpain in body wall muscle is a factor in the severity of phenotype caused by heat stress. *clp-1(tm690)* mutant worms also displayed similar levels of paralysis/death to wildtype worms after heat stress treatment, which suggests that either other *C. elegans* calpains are effecting the same phenotype after heat stress or that calpain is not the primary effector of this pathological phenotype. A greater number of worms should be assessed in future so that statistical analysis can be performed on the dataset, and video monitoring and locomotion software could be used to increase the accuracy of determining locomotion defects.

To understand the role of calpains in body wall muscle an unsuccessful attempt was made to construct a *clp-1;clp-4* double mutant because CLP-1 and CLP-4 are the only *C. elegans* proteins that have been found to be expressed in body wall muscle (5). sgRNA clones were designed and constructed that targeted deletion of a region of *clp-4* that encodes catalytic triad. 28 worms were injected and 50 worms displaying the CRISPR-Cas9 co-conversion marker were screened but a mutant was not observed through PCR screening. It was observed that CRISPR-Cas9 had

occurred in injected worms due to dumpy and roller phenotype, however no mutant bands indicated no successful events. This might be due to rarity of successful events or the sgRNA not being optimised. In future it would be worth repeating the microinjections and also screening a greater number of worms. It might also be beneficial to repeat the microinjections with a different pair of sgRNAs still targeting a deletion of the catalytic triad of *clp-4*.

4.3.1 Effects of Heat stress on *clp-1* mutants and generating a *clp-1;clp-4* double mutant using CRISPR-cas9: future work

A *clp-1;clp-4* double mutant would be a useful reagent to test physiological functional of calpains in the *C. elegans* body wall muscle and could also be used for crosses with other calpain mutants or RNAi experiments, as RNAi efficiency is reduced when used for multiple genes (142). For further physiological analysis it would also be of interest to conduct Life span assays of *clp-1(tm690), clp-1(tm690);clp-4(lf)* mutants and worms treated with calpain inhibitor, and determine the relationship between calpain activity and senescence.

Appendix

Oligonucleotides

Primer no.	Primer name	Sequence 5' to 3'
PK1241	pLN1 pCFJ_R	GCCCACTAGTGAGTCGTATTATAAG
PK1242	pLN1 pCFJ_F	AGATCTTCGAATGCATCGC
PK1243	pLN1 pMC7_insert_F	gactcactagtgggcCCGGAAAACACAATAATGC
PK1244	pLN1 pMC7_insert_R	gatgcattcgaagatctTCTTCTTCTAAATTCCCATAAAATC
PK1245	pCFJ151_F1	GCGGTAATACGGTTATCCAC
PK1246	pCFJ151_R1	GTGGATAACCGTATTACCGC
PK1247	pMC7_insert_F1	CGTTACTCGTTGGTAATCTGTC
PK1248	pMC7_insert_R1	GACAGATTACCAACGAGTAACG
PK1323	clp-1_R_LN	catGCTAGCCAAGGGTCC
PK1324	clp-1_domain_1_disrupt_F	cgactattcgaggatccaca
PK1325	clp-1_domain_3_disrupt_R	tgcaccatcatgagtatttgc
PK1345	LN_mRFP_F	ATGGCCTCCTCCGAGG
PK1355	mRFP_5'_revseq	CGCATGAACTCCTTGATG
PK1356	unc-54_3'_fwdseq	GCACTCATCACTCGCATC
PK1631	pMC7_rev	CATGCTAGCCAAGGGTCC
PK1632	pMC7_fwd	GCCTCCTCCGAGGACGTC
PK1633	LN6_fwd	cccttggctagcatgGCTGACGATGAGGAAGAA
PK1634	LN6_rev	gtcctcggaggaggcCAGTTCCCTGGAAAATATTG
PK1635	pBAD_rev	CCATGGTTAATTCCTCCTG
PK1636	pBAD_fwd	AACAAAAACTCATCTCAGAAG
PK1637	LN7_fwd	ggaggaattaaccatggGCTGACGATGAGGAAGAA
PK1638	LN7_rev	ctgagatgagtttttgttCAGTTCCTCCGACTCGAT
PK1666	clp-1_DIIko_fwd	cttggaccagatgttatgga
PK1667	clp-1_DIIko_rev	GAATTGTGGATCCTCGAATAG
PK1668	pBAD_rev_1	GGTTAATTCCTCCTGTTAGC
PK1669	pBAD_fwd_2	aacaggaggaattaaccATGGCTGACGATGAGGAAG
PK1670	pBAD_clp-1_fwd	tgagatgagtttttgttcCAGTTCCTCCGACTCGATAAATC
PK1671	pBAD_clp-1_rev	GAACAAAAACTCATCTCAGAAG
PK1791	unc-54 3' UTR RV (anti)	CATGTAGGGATGTTGAAGAG
PK1792	clp-1_intron_tm690_F	GGACTTTTGACCGACATTTATG
PK1806	pLN11 egl-19_1 fwd	CACACAAGCAGAAGATATTG
PK1807	pLN11 egl-19_1 rev	CAATATCTTCTGCTTGTGTG

PK1808	pLN11 egl-19_2 fwd	CACCTTCCTCATTCAAGAC
PK1809	pLN11 egl-19_2 rev	GTCTTGAATGAGGAAGGTG
PK1810	pLN9_insert_fwd	gaggacccttggctagcatgCGACTATTCGAGGATCCAC
PK1811	pLN9_insert_rev	acgtcctcggaggaggccatTGCACCATCATGAGTATTTG
PK1812	clp-4 del gRNA 1 fwd	TTCTGCCGTAAAACTTTGATTTGG
PK1813	clp-4 del gRNA 1 rev	AAACCCAAATCAAAGTTTTACGGC
PK1814	clp-4 del gRNA 2 fwd	TTCTGCGGTAGTAAAGGGATGCGG
PK1815	clp-4 del gRNA 2 rev	AAACCCGCATCCCTTTACTACCGC
PK1816	clp-4 del screening fwd	gagaaatatgtcaaaataggcc
PK1817	clp-4 del screening rev	gctctcaacacgctgaattc
PK1822	clp-4 del CRISPR-Cas9 HR template	caaaaatctcaatttacCTGCATAATTCTCCGTAAAACTTTGATTTGGTCTAGACGGTTGCTGGAAACTCTGGATCCTCAAATAGTGTCTTATTCCGGAG
PK1858	clp-4 del gRNA 1 fwd_corrected	TCTTGCCGTAAAACTTTGATTTGG
PK1859	clp-4 del gRNA 2 fwd_corrected	TCTTGCGGTAGTAAAGGGATGCGG

References

1. Sorimachi, H., Hata, S., and Ono, Y. (2011) Calpain chronicle--an enzyme family under multidisciplinary characterization. *Proc Jpn Acad Ser B Phys Biol Sci* **87**, 287-327
2. Brulé, C., Dargelos, E., Diallo, R., Listrat, A., Béchet, D., Cottin, P., and Poussard, S. (2010) Proteomic study of calpain interacting proteins during skeletal muscle aging. *Biochimie* **92**, 1923-1933
3. Ferreira, A. (2012) Calpain dysregulation in Alzheimer's disease. *ISRN Biochem* **2012**, 728571
4. Cuerrier, D., Moldoveanu, T., and Davies, P. L. (2005) Determination of peptide substrate specificity for mu-calpain by a peptide library-based approach: the importance of primed side interactions. *The Journal of biological chemistry* **280**, 40632-40641
5. Joyce, P. I., Satija, R., Chen, M., and Kuwabara, P. E. (2012) The atypical calpains: evolutionary analyses and roles in Caenorhabditis elegans cellular degeneration. *PLoS Genet* **8**, e1002602
6. Guroff, G. (1964) A NEUTRAL, CALCIUM-ACTIVATED PROTEINASE FROM THE SOLUBLE FRACTION OF RAT BRAIN. *The Journal of biological chemistry* **239**, 149-155
7. Murachi, T., Tanaka, K., Hatanaka, M., and Murakami, T. (1980) Intracellular Ca2+-dependent protease (calpain) and its high-molecular-weight endogenous inhibitor (calpastatin). *Adv Enzyme Regul* **19**, 407-424
8. Zhao, S., Liang, Z., Demko, V., Wilson, R., Johansen, W., Olsen, O. A., and Shalchian-Tabrizi, K. (2012) Massive expansion of the calpain gene family in unicellular eukaryotes. *BMC Evol Biol* **12**, 193
9. Ono, Y., and Sorimachi, H. (2012) Calpains: an elaborate proteolytic system. *Biochim Biophys Acta* **1824**, 224-236
10. Goll, D. E., Thompson, V. F., Li, H., Wei, W., and Cong, J. (2003) The calpain system. *Physiol Rev* **83**, 731-801
11. Richard, I., Broux, O., Allamand, V., Fougerousse, F., Chiannilkulchai, N., Bourg, N., Brenguier, L., Devaud, C., Pasturaud, P., and Roudaut, C. (1995) Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. *Cell* **81**, 27-40
12. Maki, M., Maemoto, Y., Osako, Y., and Shibata, H. (2012) Evolutionary and physical linkage between calpains and penta-EF-hand Ca2+-binding proteins. *FEBS J* **279**, 1414-1421
13. Rawlings, N. D. (2015) Bacterial calpains and the evolution of the calpain (C2) family of peptidases. *Biol Direct* **10**, 66
14. Cong, J., Goll, D. E., Peterson, A. M., and Kapprell, H. P. (1989) The role of autolysis in activity of the Ca2+-dependent proteinases (mu-calpain and m-calpain). *J Biol Chem* **264**, 10096-10103
15. Farkas, A., Tompa, P., and Friedrich, P. (2003) Revisiting ubiquity and tissue specificity of human calpains. *Biol Chem* **384**, 945-949
16. Hanna, R. A., Campbell, R. L., and Davies, P. L. (2008) Calcium-bound structure of calpain and its mechanism of inhibition by calpastatin. *Nature* **456**, 409-412
17. Shinkai-Ouchi, F., Koyama, S., Ono, Y., Hata, S., Ojima, K., Shindo, M., duVerle, D., Ueno, M., Kitamura, F., Doi, N., Takigawa, I., Mamitsuka, H., and Sorimachi, H. (2016) Predictions of Cleavability of Calpain Proteolysis by

Quantitative Structure-Activity Relationship Analysis Using Newly Determined Cleavage Sites and Catalytic Efficiencies of an Oligopeptide Array. *Mol Cell Proteomics* **15**, 1262-1280
18. Hosfield, C. M., Elce, J. S., Davies, P. L., and Jia, Z. (1999) Crystal structure of calpain reveals the structural basis for Ca(2+)-dependent protease activity and a novel mode of enzyme activation. *EMBO J* **18**, 6880-6889
19. Strobl, S., Fernandez-Catalan, C., Braun, M., Huber, R., Masumoto, H., Nakagawa, K., Irie, A., Sorimachi, H., Bourenkow, G., Bartunik, H., Suzuki, K., and Bode, W. (2000) The crystal structure of calcium-free human m-calpain suggests an electrostatic switch mechanism for activation by calcium. *Proc Natl Acad Sci U S A* **97**, 588-592
20. Campbell, R. L., and Davies, P. L. (2012) Structure-function relationships in calpains. *Biochem J* **447**, 335-351
21. Zimmerman, U. J., and Schlaepfer, W. W. (1991) Two-stage autolysis of the catalytic subunit initiates activation of calpain I. *Biochim Biophys Acta* **1080**, 275
22. Polgár, L. (2005) The catalytic triad of serine peptidases. *Cell Mol Life Sci* **62**, 2161-2172
23. Moldoveanu, T., Hosfield, C. M., Lim, D., Elce, J. S., Jia, Z., and Davies, P. L. (2002) A Ca(2+) switch aligns the active site of calpain. *Cell* **108**, 649-660
24. Moldoveanu, T., Jia, Z., and Davies, P. L. (2004) Calpain activation by cooperative Ca2+ binding at two non-EF-hand sites. *The Journal of biological chemistry* **279**, 6106-6114
25. Tompa, P., Emori, Y., Sorimachi, H., Suzuki, K., and Friedrich, P. (2001) Domain III of calpain is a ca2+-regulated phospholipid-binding domain. *Biochem Biophys Res Commun* **280**, 1333-1339
26. Fernández-Montalván, A., Assfalg-Machleidt, I., Pfeiler, D., Fritz, H., Jochum, M., and Machleidt, W. (2006) Mu-calpain binds to lipid bilayers via the exposed hydrophobic surface of its Ca2+-activated conformation. *Biol Chem* **387**, 617-627
27. Blanchard, H., Grochulski, P., Li, Y., Arthur, J. S., Davies, P. L., Elce, J. S., and Cygler, M. (1997) Structure of a calpain Ca(2+)-binding domain reveals a novel EF-hand and Ca(2+)-induced conformational changes. *Nat Struct Biol* **4**, 532-538
28. Imajoh, S., Kawasaki, H., and Suzuki, K. (1987) The COOH-terminal E-F hand structure of calcium-activated neutral protease (CANP) is important for the association of subunits and resulting proteolytic activity. *J Biochem* **101**, 447-452
29. Sorimachi, H., and Ono, Y. (2012) Regulation and physiological roles of the calpain system in muscular disorders. *Cardiovasc Res* **96**, 11-22
30. Sorimachi, H., Hata, S., and Ono, Y. (2011) Impact of genetic insights into calpain biology. *J Biochem* **150**, 23-37
31. Franco, S. J., and Huttenlocher, A. (2005) Regulating cell migration: calpains make the cut. *J Cell Sci* **118**, 3829-3838
32. Sedarous, M., Keramaris, E., O'Hare, M., Melloni, E., Slack, R. S., Elce, J. S., Greer, P. A., and Park, D. S. (2003) Calpains mediate p53 activation and neuronal death evoked by DNA damage. *The Journal of biological chemistry* **278**, 26031-26038

33. Zimmerman, U. J., Boring, L., Pak, J. H., Mukerjee, N., and Wang, K. K. (2000) The calpain small subunit gene is essential: its inactivation results in embryonic lethality. *IUBMB Life* **50**, 63-68
34. Azam, M., Andrabi, S. S., Sahr, K. E., Kamath, L., Kuliopulos, A., and Chishti, A. H. (2001) Disruption of the mouse mu-calpain gene reveals an essential role in platelet function. *Mol Cell Biol* **21**, 2213-2220
35. Dutt, P., Croall, D. E., Arthur, J. S., Veyra, T. D., Williams, K., Elce, J. S., and Greer, P. A. (2006) m-Calpain is required for preimplantation embryonic development in mice. *BMC Dev Biol* **6**, 3
36. Richard, I., Roudaut, C., Marchand, S., Baghdiguian, S., Herasse, M., Stockholm, D., Ono, Y., Suel, L., Bourg, N., Sorimachi, H., Lefranc, G., Fardeau, M., Sébille, A., and Beckmann, J. S. (2000) Loss of calpain 3 proteolytic activity leads to muscular dystrophy and to apoptosis-associated IkappaBalpha/nuclear factor kappaB pathway perturbation in mice. *J Cell Biol* **151**, 1583-1590
37. Wendt, A., Thompson, V. F., and Goll, D. E. (2004) Interaction of calpastatin with calpain: a review. *Biol Chem* **385**, 465-472
38. Csizmók, V., Bokor, M., Bánki, P., Klement, E., Medzihradszky, K. F., Friedrich, P., Tompa, K., and Tompa, P. (2005) Primary contact sites in intrinsically unstructured proteins: the case of calpastatin and microtubule-associated protein 2. *Biochemistry* **44**, 3955-3964
39. Chakraborti, S., Alam, M. N., Paik, D., Shaikh, S., and Chakraborti, T. (2012) Implications of calpains in health and diseases. *Indian J Biochem Biophys* **49**, 316-328
40. Saez, M. E., Ramirez-Lorca, R., Moron, F. J., and Ruiz, A. (2006) The therapeutic potential of the calpain family: new aspects. *Drug Discov Today* **11**, 917-923
41. Ono, Y., Saido, T. C., and Sorimachi, H. (2016) Calpain research for drug discovery: challenges and potential. *Nat Rev Drug Discov* **15**, 854-876
42. Carragher, N. O. (2006) Calpain inhibition: a therapeutic strategy targeting multiple disease states. *Curr Pharm Des* **12**, 615-638
43. Huttenlocher, A., Palecek, S. P., Lu, Q., Zhang, W., Mellgren, R. L., Lauffenburger, D. A., Ginsberg, M. H., and Horwitz, A. F. (1997) Regulation of cell migration by the calcium-dependent protease calpain. *The Journal of biological chemistry* **272**, 32719-32722
44. Dourdin, N., Bhatt, A. K., Dutt, P., Greer, P. A., Arthur, J. S., Elce, J. S., and Huttenlocher, A. (2001) Reduced cell migration and disruption of the actin cytoskeleton in calpain-deficient embryonic fibroblasts. *The Journal of biological chemistry* **276**, 48382-48388
45. Saraiva, N., Prole, D. L., Carrara, G., Johnson, B. F., Taylor, C. W., Parsons, M., and Smith, G. L. (2013) hGAAP promotes cell adhesion and migration via the stimulation of store-operated Ca2+ entry and calpain 2. *J Cell Biol* **202**, 699-713
46. Yousefi, S., Perozzo, R., Schmid, I., Ziemiecki, A., Schaffner, T., Scapozza, L., Brunner, T., and Simon, H. U. (2006) Calpain-mediated cleavage of Atg5 switches autophagy to apoptosis. *Nat Cell Biol* **8**, 1124-1132
47. Fan, Y. J., and Zong, W. X. (2013) The cellular decision between apoptosis and autophagy. *Chin J Cancer* **32**, 121-129
48. Russo, R., Berliocchi, L., Adornetto, A., Varano, G. P., Cavaliere, F., Nucci, C., Rotiroti, D., Morrone, L. A., Bagetta, G., and Corasaniti, M. T. (2011)

Calpain-mediated cleavage of Beclin-1 and autophagy deregulation following retinal ischemic injury in vivo. *Cell Death Dis* **2**, e144
49. Demarchi, F., Bertoli, C., Copetti, T., Tanida, I., Brancolini, C., Eskelinen, E. L., and Schneider, C. (2006) Calpain is required for macroautophagy in mammalian cells. *J Cell Biol* **175**, 595-605
50. Yamashima, T. (2004) Ca2+-dependent proteases in ischemic neuronal death: a conserved 'calpain-cathepsin cascade' from nematodes to primates. *Cell Calcium* **36**, 285-293
51. Syntichaki, P., Xu, K., Driscoll, M., and Tavernarakis, N. (2002) Specific aspartyl and calpain proteases are required for neurodegeneration in C. elegans. *Nature* **419**, 939-944
52. Vallejo-Illarramendi, A., Toral-Ojeda, I., Aldanondo, G., and López de Munain, A. (2014) Dysregulation of calcium homeostasis in muscular dystrophies. *Expert Rev Mol Med* **16**, e16
53. Ojima, K., Kawabata, Y., Nakao, H., Nakao, K., Doi, N., Kitamura, F., Ono, Y., Hata, S., Suzuki, H., Kawahara, H., Bogomolovas, J., Witt, C., Ottenheijm, C., Labeit, S., Granzier, H., Toyama-Sorimachi, N., Sorimachi, M., Suzuki, K., Maeda, T., Abe, K., Aiba, A., and Sorimachi, H. (2010) Dynamic distribution of muscle-specific calpain in mice has a key role in physical-stress adaptation and is impaired in muscular dystrophy. *J Clin Invest* **120**, 2672-2683
54. Huebsch, K. A., Kudryashova, E., Wooley, C. M., Sher, R. B., Seburn, K. L., Spencer, M. J., and Cox, G. A. (2005) Mdm muscular dystrophy: interactions with calpain 3 and a novel functional role for titin's N2A domain. *Hum Mol Genet* **14**, 2801-2811
55. Piva, R., Belardo, G., and Santoro, M. G. (2006) NF-kappaB: a stress-regulated switch for cell survival. *Antioxid Redox Signal* **8**, 478-486
56. Yoshikawa, Y., Mukai, H., Hino, F., Asada, K., and Kato, I. (2000) Isolation of two novel genes, down-regulated in gastric cancer. *Jpn J Cancer Res* **91**, 459-463
57. Medeiros, R., Kitazawa, M., Chabrier, M. A., Cheng, D., Baglietto-Vargas, D., Kling, A., Moeller, A., Green, K. N., and LaFerla, F. M. (2012) Calpain inhibitor A-705253 mitigates Alzheimer's disease-like pathology and cognitive decline in aged 3xTgAD mice. *Am J Pathol* **181**, 616-625
58. Ferreira, A., and Bigio, E. H. (2011) Calpain-mediated tau cleavage: a mechanism leading to neurodegeneration shared by multiple tauopathies. *Mol Med* **17**, 676-685
59. Patzke, H., and Tsai, L. H. (2002) Calpain-mediated cleavage of the cyclin-dependent kinase-5 activator p39 to p29. *The Journal of biological chemistry* **277**, 8054-8060
60. Mushtaq, G., Greig, N. H., Anwar, F., Al-Abbasi, F. A., Zamzami, M. A., Al-Talhi, H. A., and Kamal, M. A. (2016) Neuroprotective Mechanisms Mediated by CDK5 Inhibition. *Curr Pharm Des* **22**, 527-534
61. Tang, D., Borchman, D., Yappert, M. C., Vrensen, G. F., and Rasi, V. (2003) Influence of age, diabetes, and cataract on calcium, lipid-calcium, and protein-calcium relationships in human lenses. *Invest Ophthalmol Vis Sci* **44**, 2059-2066
62. Shearer, T. R., Shih, M., Azuma, M., and David, L. L. (1995) Precipitation of crystallins from young rat lens by endogenous calpain. *Exp Eye Res* **61**, 141-150

63. Paul, D. S., Harmon, A. W., Winston, C. P., and Patel, Y. M. (2003) Calpain facilitates GLUT4 vesicle translocation during insulin-stimulated glucose uptake in adipocytes. *Biochem J* **376**, 625-632
64. Brenner, S. (1974) The genetics of Caenorhabditis elegans. *Genetics* **77**, 71-94
65. White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1986) The structure of the nervous system of the nematode Caenorhabditis elegans. *Philos Trans R Soc Lond B Biol Sci* **314**, 1-340
66. Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983) The embryonic cell lineage of the nematode Caenorhabditis elegans. *Dev Biol* **100**, 64-119
67. Consortium, C. e. S. (1998) Genome sequence of the nematode C. elegans: a platform for investigating biology. *Science* **282**, 2012-2018
68. Maher, K. N., Catanese, M., and Chase, D. L. (2013) Large-scale gene knockdown in C. elegans using dsRNA feeding libraries to generate robust loss-of-function phenotypes. *J Vis Exp*, e50693
69. Moresco, J. J., Carvalho, P. C., and Yates, J. R. (2010) Identifying components of protein complexes in C. elegans using co-immunoprecipitation and mass spectrometry. *J Proteomics* **73**, 2198-2204
70. Shaye, D. D., and Greenwald, I. (2011) OrthoList: a compendium of C. elegans genes with human orthologs. *PLoS One* **6**, e20085
71. Alexander, A. G., Marfil, V., and Li, C. (2014) Use of Caenorhabditis elegans as a model to study Alzheimer's disease and other neurodegenerative diseases. *Front Genet* **5**, 279
72. Markaki, M., and Tavernarakis, N. (2010) Modeling human diseases in Caenorhabditis elegans. *Biotechnol J* **5**, 1261-1276
73. Benian, G. M., and Epstein, H. F. (2011) Caenorhabditis elegans muscle: a genetic and molecular model for protein interactions in the heart. *Circ Res* **109**, 1082-1095
74. Riddle, D. L. (1997) *C. elegans II*, Cold Spring Harbor Laboratory Press, New York
75. Francis, G. R., and Waterston, R. H. (1985) Muscle organization in Caenorhabditis elegans: localization of proteins implicated in thin filament attachment and I-band organization. *J Cell Biol* **101**, 1532-1549
76. Moerman, D. G., and Williams, B. D. (2006) Sarcomere assembly in C. elegans muscle. *WormBook*, 1-16
77. Etheridge, T., Oczypok, E. A., Lehmann, S., Fields, B. D., Shephard, F., Jacobson, L. A., and Szewczyk, N. J. (2012) Calpains mediate integrin attachment complex maintenance of adult muscle in Caenorhabditis elegans. *PLoS Genet* **8**, e1002471
78. Sokol, S. B., and Kuwabara, P. E. (2000) Proteolysis in Caenorhabditis elegans sex determination: cleavage of TRA-2A by TRA-3. *Genes Dev* **14**, 901-906
79. Yook, K., Harris, T. W., Bieri, T., Cabunoc, A., Chan, J., Chen, W. J., Davis, P., de la Cruz, N., Duong, A., Fang, R., Ganesan, U., Grove, C., Howe, K., Kadam, S., Kishore, R., Lee, R., Li, Y., Muller, H. M., Nakamura, C., Nash, B., Ozersky, P., Paulini, M., Raciti, D., Rangarajan, A., Schindelman, G., Shi, X., Schwarz, E. M., Ann Tuli, M., Van Auken, K., Wang, D., Wang, X., Williams, G., Hodgkin, J., Berriman, M., Durbin, R., Kersey, P., Spieth, J., Stein, L., and

Sternberg, P. W. (2012) WormBase 2012: more genomes, more data, new website. *Nucleic Acids Res* **40**, D735-741
80. Jospin, M., Jacquemond, V., Mariol, M. C., Ségalat, L., and Allard, B. (2002) The L-type voltage-dependent Ca2+ channel EGL-19 controls body wall muscle function in Caenorhabditis elegans. *J Cell Biol* **159**, 337-348
81. Praitis, V., Simske, J., Kniss, S., Mandt, R., Imlay, L., Feddersen, C., Miller, M. B., Mushi, J., Liszewski, W., Weinstein, R., Chakravorty, A., Ha, D. G., Schacht Farrell, A., Sullivan-Wilson, A., and Stock, T. (2013) The secretory pathway calcium ATPase PMR-1/SPCA1 has essential roles in cell migration during Caenorhabditis elegans embryonic development. *PLoS Genet* **9**, e1003506
82. Hagedorn, E. J., Yashiro, H., Ziel, J. W., Ihara, S., Wang, Z., and Sherwood, D. R. (2009) Integrin acts upstream of netrin signaling to regulate formation of the anchor cell's invasive membrane in C. elegans. *Dev Cell* **17**, 187-198
83. al, M. e. (2004) Cold Spring Harbor Symposia on Quantitative Biology 68. in ***WBPaper00006525***
84. Chen, M. (2009) *Atypical calpain and muscle function in Caenorhabditis elegans.* MSc, University of Bristol
85. Frøkjaer-Jensen, C., Davis, M. W., Hopkins, C. E., Newman, B. J., Thummel, J. M., Olesen, S. P., Grunnet, M., and Jorgensen, E. M. (2008) Single-copy insertion of transgenes in Caenorhabditis elegans. *Nat Genet* **40**, 1375-1383
86. Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991) Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. *EMBO J* **10**, 3959-3970
87. Sambrook, J. (2001) Molecular cloning : a laboratory manual. (Russell, D. W. ed., 3rd ed. Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. :
88. Engebrecht, J., and Brent, R. (2001) Preparation of plasmid DNA. *Curr Protoc Protein Sci* **Appendix 4**, Appendix 4C
89. Cohen, S. N., Chang, A. C., and Hsu, L. (1972) Nonchromosomal antibiotic resistance in bacteria: genetic transformation of Escherichia coli by R-factor DNA. *Proc Natl Acad Sci U S A* **69**, 2110-2114
90. (2017) NEBuilder. New England Biolabs
91. Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343-345
92. Seidman, C. E., and Struhl, K. (2001) Introduction of plasmid DNA into cells. *Curr Protoc Protein Sci* **Appendix 4**, 4D
93. Smith, B. J. (1984) SDS Polyacrylamide Gel Electrophoresis of Proteins. *Methods Mol Biol* **1**, 41-55
94. Towbin, H., Staehelin, T., and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**, 4350-4354
95. Momma, K., Homma, T., Isaka, R., Sudevan, S., and Higashitani, A. (2017) Heat-Induced Calcium Leakage Causes Mitochondrial Damage in Caenorhabditis elegans Body-Wall Muscles. *Genetics* **206**, 1985-1994
96. (ed.), E., and C., T. (2006) *WormBook*, WormBook
97. Lewis, J. A., and Fleming, J. T. (1995) Basic culture methods. *Methods Cell Biol* **48**, 3-29

98. Shaham (ed.), S. (January 12, 2006) Methods in cell biology. in *The C. elegans Research Community* (Shaham (ed.), S. ed., WormBook, WormBook
99. Hope, I. A. (1999) *C. elegans : a practical approach*, Oxford University Press, Oxford
100. Rothbauer, U., Zolghadr, K., Muyldermans, S., Schepers, A., Cardoso, M. C., and Leonhardt, H. (2008) A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. *Mol Cell Proteomics* **7**, 282-289
101. Pfaff, M., Du, X., and Ginsberg, M. H. (1999) Calpain cleavage of integrin beta cytoplasmic domains. *FEBS Lett* **460**, 17-22
102. Serrano, K., and Devine, D. V. (2004) Vinculin is proteolyzed by calpain during platelet aggregation: 95 kDa cleavage fragment associates with the platelet cytoskeleton. *Cell Motil Cytoskeleton* **58**, 242-252
103. Baron, C. P., Jacobsen, S., and Purslow, P. P. (2004) Cleavage of desmin by cysteine proteases: Calpains and cathepsin B. *Meat Sci* **68**, 447-456
104. MacLeod, A. R., Karn, J., and Brenner, S. (1981) Molecular analysis of the unc-54 myosin heavy-chain gene of Caenorhabditis elegans. *Nature* **291**, 386-390
105. Yang, L., Zhang, H., and Bruce, J. E. (2009) Optimizing the detergent concentration conditions for immunoprecipitation (IP) coupled with LC-MS/MS identification of interacting proteins. *Analyst* **134**, 755-762
106. Zanin, E., Dumont, J., Gassmann, R., Cheeseman, I., Maddox, P., Bahmanyar, S., Carvalho, A., Niessen, S., Yates, J. R., Oegema, K., and Desai, A. (2011) Affinity purification of protein complexes in C. elegans. *Methods Cell Biol* **106**, 289-322
107. Heesom, K. (2018) Bristol University Biomedical Sciences Proteomics Facility.
108. McIlwain, S., Mathews, M., Bereman, M. S., Rubel, E. W., MacCoss, M. J., and Noble, W. S. (2012) Estimating relative abundances of proteins from shotgun proteomics data. *BMC Bioinformatics* **13**, 308
109. Kim, Y., Kim, T.-H., and Ergün, T. (2015) The instability of the Pearson correlation coefficient in the presence of coincidental outliers. *Finance Research Letters* **13**, 243-257
110. Fischer, M., Zilkenat, S., Gerlach, R. G., Wagner, S., and Renard, B. Y. (2014) Pre- and post-processing workflow for affinity purification mass spectrometry data. *J Proteome Res* **13**, 2239-2249
111. Teo, G., Liu, G., Zhang, J., Nesvizhskii, A. I., Gingras, A. C., and Choi, H. (2014) SAINTexpress: improvements and additional features in Significance Analysis of INTeractome software. *J Proteomics* **100**, 37-43
112. Coordinators, N. R. (2018) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **46**, D8-D13
113. Angeles-Albores, D., Lee, R. Y., Chan, J., and Sternberg, P. W. (2017) Phenotype and gene ontology enrichment as guides for disease modeling in C. elegans. *bioRxiv*
114. Santella, L., Kyozuka, K., Hoving, S., Munchbach, M., Quadroni, M., Dainese, P., Zamparelli, C., James, P., and Carafoli, E. (2000) Breakdown of cytoskeletal proteins during meiosis of starfish oocytes and proteolysis induced by calpain. *Exp Cell Res* **259**, 117-126
115. Bédard, N., Jammoul, S., Moore, T., Wykes, L., Hallauer, P. L., Hastings, K. E., Stretch, C., Baracos, V., Chevalier, S., Plourde, M., Coyne, E., and Wing,

S. S. (2015) Inactivation of the ubiquitin-specific protease 19 deubiquitinating enzyme protects against muscle wasting. *FASEB J* **29**, 3889-3898
116. Molyneux, S. L., Young, J. M., Florkowski, C. M., Lever, M., and George, P. M. (2008) Coenzyme Q10: is there a clinical role and a case for measurement? *Clin Biochem Rev* **29**, 71-82
117. Spurney, C. F., Rocha, C. T., Henricson, E., Florence, J., Mayhew, J., Gorni, K., Pasquali, L., Pestronk, A., Martin, G. R., Hu, F., Nie, L., Connolly, A. M., Escolar, D. M., and Investigators, C. I. N. R. G. (2011) CINRG pilot trial of coenzyme Q10 in steroid-treated Duchenne muscular dystrophy. *Muscle Nerve* **44**, 174-178
118. Zhou, J., Ryan, A. J., Medh, J., and Mallampalli, R. K. (2003) Oxidized lipoproteins inhibit surfactant phosphatidylcholine synthesis via calpain-mediated cleavage of CTP:phosphocholine cytidylyltransferase. *The Journal of biological chemistry* **278**, 37032-37040
119. Newsom, S. A., Brozinick, J. T., Kiseljak-Vassiliades, K., Strauss, A. N., Bacon, S. D., Kerege, A. A., Bui, H. H., Sanders, P., Siddall, P., Wei, T., Thomas, M., Kuo, M. S., Nemkov, T., D'Alessandro, A., Hansen, K. C., Perreault, L., and Bergman, B. C. (2016) Skeletal muscle phosphatidylcholine and phosphatidylethanolamine are related to insulin sensitivity and respond to acute exercise in humans. *J Appl Physiol (1985)* **120**, 1355-1363
120. Lee, M. S., Kwon, Y. T., Li, M., Peng, J., Friedlander, R. M., and Tsai, L. H. (2000) Neurotoxicity induces cleavage of p35 to p25 by calpain. *Nature* **405**, 360-364
121. Rao, L., Perez, D., and White, E. (1996) Lamin proteolysis facilitates nuclear events during apoptosis. *J Cell Biol* **135**, 1441-1455
122. Worman, H. J., and Bonne, G. (2007) "Laminopathies": a wide spectrum of human diseases. *Exp Cell Res* **313**, 2121-2133
123. Korfali, N., Srsen, V., Waterfall, M., Batrakou, D. G., Pekovic, V., Hutchison, C. J., and Schirmer, E. C. (2011) A flow cytometry-based screen of nuclear envelope transmembrane proteins identifies NET4/Tmem53 as involved in stress-dependent cell cycle withdrawal. *PLoS One* **6**, e18762
124. Mercer, K. B., Miller, R. K., Tinley, T. L., Sheth, S., Qadota, H., and Benian, G. M. (2006) Caenorhabditis elegans UNC-96 is a new component of M-lines that interacts with UNC-98 and paramyosin and is required in adult muscle for assembly and/or maintenance of thick filaments. *Mol Biol Cell* **17**, 3832-3847
125. Francis, R., and Waterston, R. H. (1991) Muscle cell attachment in Caenorhabditis elegans. *J Cell Biol* **114**, 465-479
126. Gettner, S. N., Kenyon, C., and Reichardt, L. F. (1995) Characterization of beta pat-3 heterodimers, a family of essential integrin receptors in C. elegans. *J Cell Biol* **129**, 1127-1141
127. Sievers, F., and Higgins, D. G. (2018) Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci* **27**, 135-145
128. Moldoveanu, T., Gehring, K., and Green, D. R. (2008) Concerted multi-pronged attack by calpastatin to occlude the catalytic cleft of heterodimeric calpains. *Nature* **456**, 404-408
129. Momma, K., Homma, T., Isaka, R., Sudevan, S., and Higashitani, A. (2017) Heat-Induced Calcium Leakage Causes Mitochondrial Damage in. *Genetics* **206**, 1985-1994

130. Kourtis, N., Nikoletopoulou, V., and Tavernarakis, N. (2012) Small heat-shock proteins protect from heat-stroke-associated neurodegeneration. *Nature* **490**, 213-218
131. Zhang, F. e. a. M. (2015) Optimized CRISPR Design. http://crispr.mit.edu/
132. Arribere, J. A., Bell, R. T., Fu, B. X., Artiles, K. L., Hartman, P. S., and Fire, A. Z. (2014) Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in Caenorhabditis elegans. *Genetics* **198**, 837-846
133. Piehowski, P. D., Petyuk, V. A., Orton, D. J., Xie, F., Moore, R. J., Ramirez-Restrepo, M., Engel, A., Lieberman, A. P., Albin, R. L., Camp, D. G., Smith, R. D., and Myers, A. J. (2013) Sources of technical variability in quantitative LC-MS proteomics: human brain tissue sample analysis. *J Proteome Res* **12**, 2128-2137
134. Caberoy, N. B., Alvarado, G., and Li, W. (2011) Identification of calpain substrates by ORF phage display. *Molecules* **16**, 1739-1748
135. Jiang, L. Q., Wen, S. J., Wang, H. Y., and Chen, L. Y. (2002) Screening the proteins that interact with calpain in a human heart cDNA library using a yeast two-hybrid system. *Hypertens Res* **25**, 647-652
136. Joyce, P. I. (2008) *Characterisation of the atypical calpain family of C. elegans* PhD, University of Bristol
137. Kalderimis, A., Lyne, R., Butano, D., Contrino, S., Lyne, M., Heimbach, J., Hu, F., Smith, R., Stěpán, R., Sullivan, J., and Micklem, G. (2014) InterMine: extensive web services for modern biology. *Nucleic Acids Res* **42**, W468-472
138. Müller, H. M., Van Auken, K. M., Li, Y., and Sternberg, P. W. (2018) Textpresso Central: a customizable platform for searching, text mining, viewing, and curating biomedical literature. *BMC Bioinformatics* **19**, 94
139. Meissner, B., Rogalski, T., Viveiros, R., Warner, A., Plastino, L., Lorch, A., Granger, L., Segalat, L., and Moerman, D. G. (2011) Determining the sub-cellular localization of proteins within Caenorhabditis elegans body wall muscle. *PLoS One* **6**, e19937
140. Gil-Parrado, S., Popp, O., Knoch, T. A., Zahler, S., Bestvater, F., Felgenträger, M., Holloschi, A., Fernández-Montalván, A., Auerswald, E. A., Fritz, H., Fuentes-Prior, P., Machleidt, W., and Spiess, E. (2003) Subcellular localization and in vivo subunit interactions of ubiquitous mu-calpain. *J Biol Chem* **278**, 16336-16346
141. Chimura, T., Launey, T., and Yoshida, N. (2015) Calpain-Mediated Degradation of Drebrin by Excitotoxicity In vitro and In vivo. *PLoS One* **10**, e0125119
142. Gönczy, P., Echeverri, C., Oegema, K., Coulson, A., Jones, S. J., Copley, R. R., Duperon, J., Oegema, J., Brehm, M., Cassin, E., Hannak, E., Kirkham, M., Pichler, S., Flohrs, K., Goessen, A., Leidel, S., Alleaume, A. M., Martin, C., Ozlü, N., Bork, P., and Hyman, A. A. (2000) Functional genomic analysis of cell division in C. elegans using RNAi of genes on chromosome III. *Nature* **408**, 331-336